

University of Groningen

## Editing of the proteolytic system of *Lactococcus lactis* increases its bioactive potential

Huang, Chenxi; Kok, Jan

*Published in:*  
Applied and environmental microbiology

*DOI:*  
[10.1128/AEM.01319-20](https://doi.org/10.1128/AEM.01319-20)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Version created as part of publication process; publisher's layout; not normally made publicly available

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Huang, C., & Kok, J. (2020). Editing of the proteolytic system of *Lactococcus lactis* increases its bioactive potential. *Applied and environmental microbiology*, 86(18), [e01319]. <https://doi.org/10.1128/AEM.01319-20>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

1    **Editing of the proteolytic system of *Lactococcus lactis* increases its**  
2    **bioactive potential**

3    Chenxi Huang, Jan Kok\*

4

5    Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology

6    Institute, University of Groningen, Groningen, The Netherlands

7

8    Keywords: *Lactococcus lactis*; Intracellular peptidomics; Proteolytic system; Bioactive peptides

9

10    For correspondence. Email jan.kok@rug.nl; Tel. +31 50 363 2111

11    Running title: Editing *L. lactis* bioactive potential

12

13

## 14    **Abstract**

15    Large-scale mass spectrometry-based peptidomics for bioactive peptide discovery is  
16    relatively unexplored because of challenges in intracellular peptide extraction and small  
17    peptide identification. Here we present an analytical pipeline for large-scale intracellular  
18    peptidomics of *Lactococcus lactis*. It entails an optimized sample preparation protocol for *L.*  
19    *lactis*, used as an “enzyme complex” to digest  $\beta$ -casein, an extraction method for its  
20    intracellular peptidome, and a peptidomics data analysis and visualization procedure. In  
21    addition, we proofread the publicly available bioactive peptide databases and obtained an  
22    optimized database of bioactive peptides derivable from bovine  $\beta$ -casein. We used the  
23    pipeline to examine cultures of *L. lactis* MG1363 and a set of 6 isogenic multiple peptidase  
24    mutants incubated with  $\beta$ -casein. We observed a clearly strain-dependent accumulation of  
25    peptides with several bioactivities, such as ACE inhibitory, DPP-IV inhibitory,  
26    immunoregulatory functions. The results suggest that both the number of different bioactive  
27    peptides and the bioactivity diversity can be increased by editing the proteolytic system of *L.*  
28    *lactis*. This comprehensive pipeline offers a model for bioactive peptide discovery in  
29    combination with other proteins and might be applicable to other bacteria.

## 30    **Importance**

31    Lactic acid bacteria (LAB) are very important for the production of safe and healthy human  
32    and animal fermented foods and feed and, increasingly more, in the functional food industry.  
33    The intracellular peptidomes of LAB are promising reservoirs of bioactive peptides. We show  
34    here that targeted genetic engineering of the peptide degradation pathway allows steering

35 the composition of the peptide pool of the LAB *Lactococcus lactis* and producing peptides  
36 with interesting bioactivities. Our work could be used as a guideline for modifying proteolytic  
37 systems in other LAB to further explore their potential as cell peptide factories.

38

## 39 Introduction

40 Members of the diverse group of lactic acid bacteria (LAB) have been associated with food  
41 fermentations since ancient times. LAB have been used as starter cultures for dairy  
42 production more than 100 years ago, which has given them their current industrial and  
43 economic importance (1). However, the value of those fermented products such as yogurt or  
44 cheese have remained restricted to providing basic nutrition, flavor, and texture. The  
45 potential of LAB as production organisms for functional foods is still largely unexplored.  
46 *Lactococcus lactis*, for instance, depends on the milk protein casein as the major source of  
47 nitrogen since it is auxotrophic for several amino acids. Casein degradation is accomplished in  
48 a three-step process. First, the extracellular cell envelope-associated proteinase (PrpP)  
49 degrades casein into oligopeptides. Second, (a selection of) these peptides are internalized  
50 via an oligopeptide transport system (Opp), while in the third and last step, multiple  
51 peptidases hydrolyze the oligopeptides into smaller peptides and, ultimately, into amino  
52 acids that are then available for *de novo* protein synthesis and other metabolic activities (2).

53 Functional food for health promotion or disease risk reduction has attracted the interest of  
54 food industries worldwide, with dairy products as one of the most popular categories. Milk  
55 proteins encode bioactive peptide sequences, which can be released by hydrolysis (3). These  
56 short peptides can display a spectrum of biological functions such as angiotensin-converting  
57 enzyme (ACE) inhibitory, dipeptidyl peptidase 4 inhibitory (DPP-IV-I), immunoregulatory,  
58 antioxidant, antimicrobial, and opioid activities. Bioactive peptides from milk proteins can be  
59 obtained via the action of microbial or non-microbial enzymes. Milk fermentation processes  
60 executed by LAB are preferable ways to release these peptides because of the food grade

61 safety status of these organisms. Moreover, proteolytic systems of LAB, especially that of *L.*  
62 *lactis*, have been comprehensively studied with respect to the genes and enzymes involved  
63 and their regulation (4, 5).

64 A lot of research has focused on the production of milk-derived bioactive peptides using LAB.  
65 Two major ways of bioactive peptide discovery can be discerned. First, casein proteins are  
66 either digested by purified digestive enzyme (trypsin) or LAB proteinase(s), after which the  
67 obtained products are identified (6–8). Second, an LAB cell culture is mixed with milk proteins  
68 and the supernatant is subsequently further characterized (9–11). These studies have  
69 identified numerous casein-derived bioactive peptides, most of which having ACE-inhibitory  
70 activity. From an application point of view, the costs of employing purified enzymes are too  
71 high for industrial-scale use. On the other hand, only utilizing the culture supernatant of  
72 proteolytically active cells does not exploit the full potential of the LAB as in that case only the  
73 proteinase specificity is being utilized while the activity of the more than 10 intracellular  
74 peptidases and possible hidden intracellular bioactive peptides are being ignored.

75 Our understanding of the intracellular peptide pool in LAB during growth in a milk medium,  
76 and the possible presence of bioactive variants, is limited to nearly absent due to the  
77 technical obstacles of preparing and separating the complex samples and the subsequent  
78 identification of the small peptides (12). Recent rapid developments in nanoscale liquid  
79 chromatography coupled to tandem mass spectrometry (nano LC-MS/MS) technology and in  
80 algorithms for peptide identification have resulted in a dramatic increase in proteomics  
81 research and its sub-field, peptidomics (13, 14).

82 In this study, we engineered the proteolytic system of the *L. lactis* model strain MG1363 and  
83 describe a robust and comprehensive analytical framework of cell-casein incubation  
84 conditions, intracellular peptidome extraction, data analysis and visualization, and, ultimately,  
85 identification of casein-derived bioactive peptides produced by *L. lactis* MG1363 and six of its  
86 isogenic peptidase mutants. As proof-of-concept, this work offers a pipeline for the analysis  
87 and visualization of the intracellular peptidome of bacteria and explores the possibility of  
88 applying *L. lactis* or, for that matter, other bacteria as a cell factory to produce bioactive  
89 peptides.

90

## 91 Results

### 92 Engineering of *L. lactis* proteolytic system

93 The aim of this study is to build an analytical framework for the analysis of the intracellular  
94 peptidome of *L. lactis* and to discover (putative) bioactive peptides upon degradation of  $\beta$ -  
95 casein by the organism. To kick-start  $\beta$ -casein degradation by the *L. lactis* model strain  
96 MG1363, an extracellular cell wall-anchored proteinase PrtP (caseinase) is needed. The parent  
97 strain of *L. lactis* MG1363, *L. lactis* NCDO712, carries the 55-kb PrtP proteinase and lactose  
98 plasmid pLP712. This plasmid is too large to easily reintroduce in MG1363 and its peptidase  
99 knockout derivatives while it also contains one of the oligopeptidase genes, *pepF<sub>1</sub>* (15).  
100 Therefore, a new plasmid was constructed that encodes the proteinase PrtP and its maturase  
101 PrtM (16) from pLP712, named pCH020. *L. lactis* MG1363 possesses 15 intracellular peptidases  
102 that together degrade the PrtP-liberated casein-derived oligopeptides that are internalized  
103 by the oligopeptide permease Opp. The peptidase complement will ultimately result in the  
104 decomposition of the oligopeptides into shorter peptides and free amino acids. Undigested  
105 oligopeptides and peptidase-digested shorter versions of these peptides might possess  
106 bioactivities.

107 By removing different (groups of) peptidases, more and a greater variety of intracellular  
108 peptides should accumulate, increasing the chance of discovering (novel) bioactive peptides  
109 (Fig 1A). A total of thirty-seven single and multiple isogenic peptidase mutants were  
110 constructed from *L. lactis* MG1363 (Table 1) by employing sixteen peptidase gene  
111 replacement vectors based on the replication-deficient plasmid pCS1966 (17) (Fig 1B). Several



112 multi-peptidase deletion mutants were designed based on peptidase functional groups. Thus,  
113 four mutants were obtained in which all endopeptidases (*MGΔpepOF<sub>2</sub>O<sub>2</sub>*), all  
114 aminopeptidases except *PepM* (*MGΔpepANCpcp*; see below), all proline-specific peptidases  
115 (*MGΔpepXPQ*) or the di-/tripeptidases (*MGΔpepVD<sub>A</sub>TD<sub>B</sub>*) had been removed. In addition, in a  
116 strategy to delete as much peptidase genes as possible, half of them were deleted in strains  
117 *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>* and *MGΔpepNXOTCVD<sub>A</sub>*. Note that *pepM* is an essential gene in MG1363  
118 and is thus present in all peptidase mutants.

119 Each peptidase mutant that was examined with respect to its peptidome carried the plasmid  
120 pCH020. The strains carrying pCH020 were labeled PrtP<sup>+</sup>, e.g. MG1363(PrtP<sup>+</sup>) (See Table 1), but  
121 in the presentation of the results below, the addition PrtP<sup>+</sup> is omitted for reasons of simplicity.

## 122 **Optimization of the intracellular peptidomics workflow**

123 In order to obtain high-quality LC-MS data and convincing peptide identification results, three  
124 aspects were considered: the quality of *in vivo* β-casein degradation, *L. lactis* intracellular  
125 peptidome extraction, and the peptide identification algorithm. To optimize sample  
126 preparation for LC-MS-based intracellular peptidomics, each step of the workflow was  
127 considered (Fig 2). Our previous time-series RNA-seq results (18) revealed that the proteolytic  
128 system of *L. lactis* MG1363 is relatively highly active during the log-phase of growth, thus we  
129 chose to harvest cells in the mid-log phase ( $OD_{600} \approx 1$ ) to start the *in vivo* β-casein degradation.  
130 Preliminary experiments employing different β-casein concentrations (1, 2, or 4 mg/ml) and  
131 incubation times (0.5, 1, 2, or 4h, or overnight) were tested and ultimately 4 mg/ml β-casein  
132 and 3.5 h incubation time were chosen to achieve a proper balance between sample quality

133 and time management (data not shown). Since Gram-positive bacteria such as *L. lactis* have a  
134 thick cell wall, obtaining the intracellular peptidome requires cell disruption using mechanical  
135 forces (19). Ultrasonication and the mini-beadbeater were tested, with both set-ups yielding  
136 similar results. However, when performing nanoLC-MS on the samples, it was observed that  
137 the sonicator probe introduced an overwhelming polyethylene glycol (PEG) contamination in  
138 the peptide fraction between 150-600 Da. This problem did not occur using glass-beads and  
139 the mini-beadbeater to break the cells (data not shown). Since the focus is to identify  
140 bioactive peptides, the intracellular proteome was enriched for small peptides by using the  
141 flow-through obtained after centrifuging the proteome sample over a 3-kDa cut-off filter,  
142 prior to analysis by nanoLC-MS/MS (see the M&M section). All peptidome samples obtained in  
143 this way were analyzed in biological triplicates. Excluding the  $\beta$ -casein *in vivo* degradation  
144 time, this optimized sample preparation protocol for rapid intracellular peptide extraction,  
145 from the breaking open of the cells to the filtering through the 3-kDa cut-off filter can be  
146 performed within one hour.

147 After obtaining the peptidome raw data, in order to find the most suitable search algorithm  
148 for our dataset, we tested 9 commonly used search engines for peptide identification. The  
149 PEAKS search engine (20) was used in combination with the commercial proteomics platform  
150 PEAKS studio (<http://www.bioinform.com/peaks-studio/>), while for the other 8 SearchGUI was  
151 employed, an open-source interface configuring and running proteomics searches (21). All  
152 search engines were tested under the same setting using the raw data from MG1363  
153 triplicates (see details in the M&M section). As Fig 2C shows, 5 out of 8 search engines in  
154 SearchGUI gave relatively the same level of unique peptide identifications: Tide (22) and

155 Comet (23) identified more peptides (both >400) than X! Tandem (24), MyriMatch (25), and  
156 MS-GF+ (26) (all three >300), while MS Amanda (27), OMSSA (28) and Andromeda (29) did not  
157 work properly for our data sets. Despite the fact that the MS-GF+ identification output is  
158 lower (354) it has good reproducibility since the triplicates examined with MS-GF+ are  
159 clustered in the heatmap shown in Fig 2D. As for PEAKS, it identified significantly more  
160 unique peptides (947) than the other search engines and is also reproducible (Fig 2D). Thus,  
161 we chose PEAKS as the search engine in this work. The identified peptides and proteins were  
162 further analyzed and visualized by R and the web tool Peptigram (30).

### 163 **Peptidase deletion results in different intracellular peptidomes**

164 Fig 3 and Fig 4 give a gradually more detailed account of the results obtained for the 7 strains  
165 tested, ranging from their peptidomics profiles to gene ontology analyses. Fig 3A shows the  
166 numbers of identified peptide spectrum matches (PSMs), of unique peptides, and of unique  
167 proteins in the various intracellular peptidomes. From this figure it is clear that the deletion of  
168 all endopeptidase genes significantly increases the intracellular peptide pool. As for strain  
169 *MGΔpepOF<sub>2</sub>O<sub>2</sub>*, 3.2 k PSMs were detected while 2.6 k PSMs were obtained from the sample of  
170 the parent strain MG1363. If in the strain lacking all endopeptidase activity, peptidases from  
171 other functional groups, namely the aminopeptidases N and C, the proline-specific peptidase  
172 PepX and the tri-peptidase PepT (*MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*) were removed, the number of PSMs  
173 almost doubled compared to MG1363 (5.1 k vs. 2.6 k; see Fig 3A). Also, in strain  
174 *MGΔpepNXOTCVD<sub>A</sub>* a higher number PSMs (2.9 k) was detected than in MG1363 but it was  
175 much lower than in *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*. As these two multi-peptidase mutant strains have 5  
176 peptidase gene deletions in common, it is highly likely that deletion of the endopeptidases

177 PepF<sub>2</sub> and PepO<sub>2</sub> is responsible for most of the increase in PSMs. The multi-aminopeptidase  
178 mutant strain *MGΔpepANCpcp* has slightly more PSMs detected than that of MG1363. All  
179 these observations show that by eliminating (multiple) general peptidases, the intracellular  
180 peptide pool will increase in both quantity and diversity. When peptidases with similar  
181 specificities are removed, such as in strain *MGΔpepXPQ* lacking several proline-specific  
182 peptidases, or in a strain deficient for peptidases playing important roles in the last stages of  
183 peptide degradation (di-/tripeptidases mutant *MGΔpepVDA<sub>A</sub>TD<sub>B</sub>*), a dramatic decrease in the  
184 number of unique identified peptides is seen relative to strain MG1363. Strain *MGΔpepVDA<sub>A</sub>TD<sub>B</sub>*  
185 produce less than half the PSMs of MG1363, which might be due to the fact that deletion of  
186 *pepV* affects cell wall synthesis, which ultimately disturbs other biological processes such as  
187 nitrogen metabolism (31).

188 Hierarchical clustering of the identified peptides in the triplicate samples of each strain was  
189 performed to assess the robustness of the developed methodology. The results presented in  
190 Fig 3B show that the biological replicates of each strain delivered data of good quality and  
191 reproducibility. The proteins identified in each mutant were compared with those of MG1363  
192 on the basis of their functional grouping (Fig 3C). Approximately two thirds of the proteins  
193 identified in the peptidome of *L. lactis* MG1363 can be detected in each of the peptidase  
194 mutants. Notably, *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>* and *MGΔpepNXOTCVD<sub>A</sub>* are the top two strains with  
195 respect to the number of unique proteins (168 and 97, respectively). Gene ontology (GO)  
196 enrichment analyses were performed in order to investigate the functional profile of the  
197 identified proteins from each strain and to evaluate the effects of peptidase deletions on the  
198 peptidomes of the respective *pep* mutants. The cellular function grouping of the identified

199 proteins of MG1363 and its six isogenic *pep* mutants is shown in Fig 4. For well-characterized  
200 proteins, the top three significantly enriched categories in all strains are translation (J), cell  
201 cycle control (D), and replication (J). Around one quarter to one third of the proteins are  
202 poorly characterized (R, S). Suppl. Fig 1 shows the details of the overlap in the peptidomes of  
203 all 7 strains examined. Together with Fig 3 it can be seen that, although much more unique  
204 proteins were identified in strain *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*, the total number of biological function  
205 groups did not increase.

206 **(Endo)Peptidase mutants accumulate β-casein peptides that differ in physicochemical**  
207 **properties**

208 After having analyzed the intracellular peptidome profiles for the presence of peptides  
209 derived from proteins expressed by those strains, we proceeded by examining the β-casein-  
210 derived peptides therein. These peptides and their relative intensities were visualized using  
211 the web tool Peptigram. In *L. lactis* strain MG1363, upon digestion of β-casein by the  
212 extracellular proteinase PrtP and uptake of oligopeptides by the Opp system, peptides were  
213 retrieved that cover the majority of the β-casein sequence (Fig 5A). Several regions in β-casein  
214 that are not represented or retrieved in the intracellular peptide pool are shown as gaps. The  
215 first 2 gaps represent fragment 1-15 (f1-15) and f30-40 of β-casein, which are observed in all  
216 mutants. The first gap f1-15 is the signal peptide which exist in β-casein precursor (Uniprot  
217 P02666) which contains 224 amino acids, while the β-casein (Cat #C6905, Sigma) we used  
218 contains 209 amino acid residues (f16-224). The other 7 gaps seen in the MG1363-derived  
219 peptide pattern are covered by the peptidome of one or more of the other mutant(s). It is  
220 clear that all *pep* mutants produce different β-casein peptide profiles. In the intracellular

221 peptidome of the strain lacking all endopeptidase activity, *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*, the highest  
222 relative intensity (dark green area in Fig. 5A) is seen around β-casein f180-200. This strain is  
223 also the most promising mutant with respect to possessing (more) β-casein-derived putative  
224 bioactive peptides, since the identified peptides from its intracellular peptidome cover almost  
225 all parts of the β-casein molecule and at the same time they have quite high intensities (Fig  
226 5A). It has to be noted that in mass spectrometry, peptide intensity relies on peptide  
227 ionization capacity in addition to peptide abundance and, therefore, the observed intensities  
228 cannot directly be translated to peptide concentrations. However, for the same region of β-  
229 casein, e.g. f180-200, the peptide intensities obtained with *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>* is dramatically  
230 higher than that obtained with the other strains. This implies that in the cytoplasm of this  
231 strain many more peptides from this region are present than in the cytoplasm of the other  
232 strains. Peptides identified in *MGΔpepOF<sub>2</sub>O<sub>2</sub>*, *MGΔpepNXOTCVD<sub>A</sub>*, and *MGΔpepANCPcp* also  
233 cover more of the β-casein molecule than seen in MG1363, which means that those *pep*  
234 mutants possess some β-casein-derived peptides that do not exist in the wildtype strain.  
235 Strains *MGΔpepVD<sub>A</sub>TD<sub>B</sub>* and *MGΔpepXPQ* produce significantly less PSMs and peptides (Fig 3A)  
236 and, clearly, their peptidomes also contain less β-casein-derived peptides.

237 We then examined the physicochemical properties of the peptidomes of the various strains.  
238 For the β-casein-derived peptides, those obtained with MG1363 and its *pep* mutants have  
239 very different distributions in each physicochemical property. As expected, compared to the  
240 wildtype, the two endopeptidase mutants (*MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>* and *MGΔpepOF<sub>2</sub>O<sub>2</sub>*) contained  
241 more longer β-casein-derived peptides, while shorter peptides are present in the  
242 exopeptidase mutants (see Fig 5B, Length). In agreement with its genetic make-up, the

243 mutant in which proline-specific peptidase genes have been deleted, strain MG $\Delta$ pepXPQ,  
244 produces much more proline-containing peptides than all other strains (see Fig 5B, Proline).

#### 245 **Optimization of $\beta$ -casein-derived bioactive peptide database**

246 *L. lactis* proteinase PrtP is a  $\beta$ -casein specific caseinase. To allow identifying bioactive peptides  
247 in our peptidome dataset, a comprehensive review of  $\beta$ -casein-derived bioactive peptides  
248 was performed of commonly used bioactive peptide databases. The most popular of those  
249 databases for milk-derived proteins are BIOPEP, MBPDB, and EROP-Moscow. There are pros  
250 and cons for each of these databases. BIOPEP (32) has more peptide entries but does not give  
251 proper literature references for each entry. Although EROP-Moscow (33) does have literature  
252 references, these have not been updated in recent years. Also, the query page of EROP-  
253 Moscow does not support multiple sequence searches. MBPDB (34) performs best in both  
254 searching and literature updating but it does not cover all peptides from the other two  
255 databases. For example,  $\beta$ -casein f(75-81) YPFPGPI is present and labeled in all 3 database as  
256 having opioid activity. BIOPEP does not provide a reference, while EROP-Moscow and MBPDB  
257 do. However, EROP-Moscow only provides the first research paper revealing the opioid  
258 activity (35), while there are 3 subsequent papers also proving the opioid activity of this  
259 peptide. In addition, 4 other bioactivities have been reported for this peptide over the past 2  
260 decades in MBPDB: increased satiety (36), anxiety reducing (37), anticancer (38), and ACE-  
261 inhibitory (6).

262 We combined and curated these three databases by proofreading the data for each  $\beta$ -casein-  
263 derived bioactive peptide, excluding those for which: A) no reference was provided; B) a



reference was given but the activity was hypothetical or predicted only (39); C) data was delivered but there is a discrepancy between the sequence from reference literature and the database (40); and D) only a bitter taste was recorded (41). Thus, we obtained 176 unique bioactive peptides of which, after excluding those falling under definitions A-D, 136 bioactive peptides remained (Fig 4A). They were grouped by their bioactivities; it is clear that more than half of the bioactive peptides have ACE-inhibitory activity (82/136 BPs). The second big activity group contains around 20 peptides with immunomodulatory or dipeptidyl peptidase IV inhibitory (DPP-IV-I) activity. The third, much smaller, group includes peptides with antimicrobial, opioid, antioxidation, or prolyl-endopeptidase inhibitory activities. Note that 18 of the 136 peptides possess multiple bioactivities (Fig 6B).

#### ***L. lactis* peptidase mutants produce more bioactive peptides than the wildtype strain**

Table 2 summarizes the bioactive peptides identified through nanoLC-MS/MS in the intracellular peptidomes of each strain by comparing the identified  $\beta$ -casein-derived peptides to our curated bioactive peptide database. As expected on the basis of the results presented above, *MG $\Delta$ pepNXOTCF<sub>2</sub>O<sub>2</sub>* is the most promising strain, as 18 bioactive peptides could be identified in its intracellular peptidome. Most of the peptides have ACE-inhibitory activity (11/18), while other bioactivities are also found, such as DPP-IV inhibitory, antioxidative, antimicrobial and immunoregulatory activities. Strains *MG $\Delta$ pepNXOTCVD<sub>A</sub>* and *MG $\Delta$ pepANCpcp* are second with respect to the number of bioactive peptides that are obtained with these strains. Most of the  $\beta$ -casein-derived peptides identified in both strains have the same sequence (7/9). All  $\beta$ -casein-derived bioactive peptides identified in the peptidome of *MG $\Delta$ pepNXOTCVD<sub>A</sub>* are also present in that of *MG $\Delta$ pepNXOTCF<sub>2</sub>O<sub>2</sub>*. As for



286 *MGΔpepANCpcp*, except AVYPYQR, the other 8 bioactive peptides are also observed in  
287 *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*. The β-casein-derived peptidome of *MGΔpepOF<sub>2</sub>O<sub>2</sub>*, contains 6 bioactive  
288 peptides; except for peptide VPVEPFTE, the other 5 peptides are also present in the samples  
289 of *MGΔpepOF<sub>2</sub>O<sub>2</sub>*. No bioactive peptides were observed when using *MGΔpepVD<sub>A</sub>TD<sub>B</sub>* to  
290 degrade β-casein and only 2 were found when employing strain *MGΔpepXPQ*. As explained  
291 above, these might be caused by the disruption of other biological processes such as  
292 peptidoglycan biosynthesis when deleting the dipeptidase PepV (31).

293

## 294 Discussion

295 In this study we present an analytical framework consisting of peptidome extraction, followed  
296 by tandem mass spectrometric identification and bioinformatic analysis to untangle the  
297 intracellular peptidome of *L. lactis* and to assess the potential of this organism as a cell factory  
298 for the production of bioactive peptides. The developed protocol is reproducible and can be  
299 performed in less than 2 h from peptidome extraction to mass spectrometric analysis.

300 The quality of peptide identification relies on a suitable searching algorithm. PEAKS studio  
301 identified more peptides in our data sets than all the other search engines from SearchGUI  
302 (Fig 2C) while it is also user-friendly because of its well-designed interface. However, when  
303 this commercial tool is not accessible because of its price, SearchGUI could be a good  
304 alternative because it includes the mainstream open-source search engines. MS-GF+ and  
305 Andromeda are the top algorithms in the proteomics/peptidomics field. MS-GF+ delivered  
306 relatively good and reproducible identification results. However, unlike PEAKS, it does not  
307 provide the relative intensities of identified peptides, precluding a visualization of  $\beta$ -casein  
308 digestion profiles as presented in Fig 5A. Thus, for data analysis consistency, we did not  
309 combine the results from PEAKS and MS-GF+. Andromeda, the search engine of MaxQuant, is  
310 designed for large mass-spectrometric data sets but is mostly geared towards human  
311 proteomes/peptidomes and is unsuitable for our bacterial peptidomics data. Here we focused  
312 on identifying bioactive peptides, which normally contain 2 to 20 amino acid residues (42).  
313 Note that the oligopeptide transport system (Opp) of *L. lactis* possesses the capacity to  
314 transport peptides from 4 up to at least 18 residues (43). We therefore set the mass  
315 spectrometry detection window to 170-2000 Da, which generally covers peptides containing

316 2 to 18 amino acid residues considering that the average molecular weight of an amino acid is  
317 110 Da. The gap regions in the  $\beta$ -casein peptide profile of each strain (Fig 5A) might be due to  
318 the fact that some peptides have more than 18 amino acids residues thus beyond our  
319 detection window.

320 The fact that the  $\beta$ -casein-derived bioactive peptides identified in the intracellular peptidome  
321 of each strain treated under the same conditions differs, demonstrates the potential of these  
322 *pep* mutants. The  $\beta$ -casein-derived peptides obtained with the wildtype strain MG1363 cover  
323 the whole protein (Fig 5A) with bioactive peptides only originating from  $\beta$ -casein f(70-110)  
324 (Table 2). Novel peptides are found when employing the *pep* mutant strains, which shows the  
325 potential of  $\beta$ -casein as bioactive peptides source in combination with *L. lactis* serving as an  
326 enzyme complex. The *L. lactis* peptidases have been classified in different groups on the basis  
327 of their cleavage specificity. For example, enzymes in the endopeptidase group will cleave  
328 internally in an oligopeptide and knocking them all out should result in the accumulation of  
329 relatively longer peptides. Take peptide TQTPVVPPFLQPEVM in Fig 7 as an example. It might  
330 be that in MG1363, peptide TQTPVVPPFLQPEVM was internalized and hydrolyzed into  
331 TQTPVVPPFLQPE/VM, TQTPVVPPF/LQPEVM, and TQTPVVPP/FLQPEVM. The longer N-  
332 terminal parts in each case were detected while the C-terminals were not. This may be due to  
333 the detection limit of the LC/MS equipment (VM), or the C-terminal peptides were degraded  
334 further by other intracellular peptidases (LQPEVM, FLQPEVM). In mutant  $MG\Delta pepOF_2O_2$ ,  
335 lacking the three endopeptidases, TQTPVVPPFLQPEVM would not be degraded, leading to  
336 its accumulation (Fig 7). In fact, the three peptides TQTPVVPPFLQPEVM, TPVVPPFLQPEVM,  
337 and PVVVPPFLQPEVM were only detected in the mutants  $MG\Delta pepOF_2O_2$  and

338 *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>*, which both lack all endopeptidases. In some other mutants, smaller  
339 derivative peptides identical to those in MG1363 can be observed, i.e. TQTPVVVPP and  
340 PVVVPPF were also identified in strains *MGΔpepNXOTCVD<sub>A</sub>* and *MGΔpepANCPcp*, respectively,  
341 while TPVVVPP was also identified in strain *MGΔpepXPQ*. These observations indicate that one  
342 or more of the endopeptidases prefers cutting C-terminal of glutamic acid, phenylalanine,  
343 and proline residues, which coincide with a previous study showing that PepF can (likes to)  
344 cut C-terminal proline and phenylalanine (44).

345 On the one hand, this presence/absence of certain (groups of) peptidases might liberate  
346 interesting bioactivities from the β-casein molecule and on the other hand release enough  
347 free amino acids so that *L. lactis* growth and functioning is not severely affected. When a  
348 group of specialty peptidases is removed a problem might arise if they are not only  
349 responsible for β-casein digestion but also are important in other metabolic pathway(s),  
350 disrupting certain essential processes. A clear example is the role that PepV plays in  
351 peptidoglycan synthesis (31).

352 Fermentation is an easy and cost-effective method to generate bioactive peptides in  
353 fermented milk products. This study presents a comprehensive analysis of the *L. lactis*  
354 intracellular peptidome after *in vivo* β-casein degradation. The work suggests that the  
355 number of different bioactive peptides and the bioactivity diversity can be increased by  
356 editing the proteolytic system of this LAB starter strain. *L. lactis* *MGΔpepNXOTCF<sub>2</sub>O<sub>2</sub>* has the  
357 best performance in producing peptides with high intensities of peptides that have a variety  
358 of bioactivities. It might thus potentially be useful as bioactive peptide cell factory. The fact  
359 that the peptides are intracellular should make them less sensitive to for instance digestive

enzymes. Our work could also be used as a guideline for modifying proteolytic systems in other LAB, and further analyzing and visualizing the intracellular proteome/peptidome data to explore their potential as peptide cell factories.

### Acknowledgements

We thank Dr. Anne de Jong for bioinformatics support and Dr. Hjalmar Permentier for his nanoLC/MS/MS expertise. CH was supported by a scholarship from China Scholarship Council (CSC, file nr. 201505990303).

### Materials and Methods

#### Bacterial Strains and Culture Conditions

Bacterial strains used in this study are listed in Table 1. *Lactococcus lactis* MG1363 and its derivatives were cultivated in M17 medium (Cat # DF1856-17-4, BD Difco, Detroit, MI, USA) containing 0.5% (w/v) glucose (GM17) at 30°C. Erythromycin (Cat # E6376, Sigma-Aldrich, Santa Clara, CA, USA) was added at a final concentration of 5 µg/ml when required. Chemically defined SA medium with 0.5% (w/v) glucose and 20 µg/ml 5-fluoroorotic acid (5-FOA; Cat #F5013, Sigma-Aldrich, Santa Clara, CA, USA) as a sole pyrimidine source was used for the generation of chromosomal knockouts, as described previously (17). *Escherichia coli* DH5α was used for cloning purposes; it was cultivated aerobically at 37°C in LB medium (Cat # LMM01, Formedium, Norfolk, UK) with erythromycin at a final concentration of 200 µg/ml when required.

379 **Recombinant DNA techniques and oligonucleotides**

380 Standard molecular cloning techniques were performed essentially as described (Sambrook  
381 and Russell, 2001). Chromosomal DNA from *L. lactis* was isolated using the GenElute™  
382 Bacterial Genomic DNA Kit (Cat # NA2110-1KT, Sigma-Aldrich, Santa Clara, CA, USA). Plasmids  
383 and PCR products were isolated and purified using the NucleoSpin® Plasmid EasyPure kit (Cat  
384 # MN 740727.250, Macherey-Nagel, Leiden, the Netherlands) and NucleoSpin Gel & PCR  
385 Clean-up kit (Cat # MN 740609.250, Macherey-Nagel, Leiden, the Netherlands), respectively,  
386 according to the manufacturers' instructions. PCR reactions for cloning purpose were  
387 performed with Phusion High-Fidelity DNA Polymerase (Cat # F530L, Thermo Fisher Scientific,  
388 MA, USA) according to the manufacturer's protocol; Enzymes were purchased from  
389 Fermentas (Thermo Fisher Scientific, MA, USA) and New England Biolabs (Ipswich, MA, USA).  
390 Colony PCR reactions were performed by homemade Pfu polymerase. Inserts and linearized  
391 vector were fused using the Quick-Fusion Cloning kit (Cat # B22612, BioConnect) according to  
392 the manufacturer with the modification that half of the recommended volume per reaction  
393 was used. Oligonucleotides employed in this study are listed in Supplemental Table 1 and  
394 were purchased from Biolegio BV (Nijmegen, The Netherlands). Competent *E. coli* cells were  
395 transformed using heat-shock (45), while electrocompetent *L. lactis* cells were transformed  
396 using electroporation (46) with a Bio-Rad Gene Pulser (Bio-Rad Laboratories, CA, USA). All  
397 nucleotide sequencing was performed at Macrogen Europe (Amsterdam, The Netherlands).

398 **Construction of integration plasmids for knocking out peptidase genes from *L. lactis***

399 **MG1363**

400 All plasmids that were used or constructed during this study are listed in Table 3. Relevant  
401 regions of all plasmids were sequenced to confirm their nucleotide sequences. All integration  
402 plasmids were constructed using the same workflow that is described here for only one,  
403 pCH001, as an example: Linearized vector pCS1966 was amplified using primers  
404 pCS1966\_1FW/ pCS1966\_1RV. Primers pairs pCH-0017/pCH-0018 and pCH-0019/pCH-0020  
405 were used, respectively, to obtain upstream (UP\_F2) and downstream (DOWN\_F2) regions of  
406 peptidase gene *pepF*<sub>2</sub>. Primer pair pCH-0017/pCH0020 was used to perform an overlap PCR to  
407 obtain the flanking region UP+DOWN\_F2. Primers pCH-0017 and pCH-0020 contain 15  
408 nucleotides at one end, overlapping with the sequence on the 5' end of the linearized vector,  
409 followed by the flanking region of *pepF*<sub>2</sub> gene, and 15 nucleotides overlapping with the  
410 sequence on the 3' end of the linearized vector. The fragment UP+DOWN\_F2 was fused with  
411 the linearized vector using Quick-Fusion, after which the reaction mixture was directly used to  
412 transform competent *E. coli* DH5α. The resulting vector was designated pCH001. Primers pCH-  
413 0083/0099/0100 were used for colony PCR and nucleotide sequencing confirmation.

414 **Construction of *L. lactis* (multi)peptidase knockout mutants**

415 All peptidase gene knockout strains were made using (a repetition of) the same workflow that  
416 is only described here for the construction of the multiple endopeptidase mutant strain  
417 CH018 as an example: Integration plasmid pCH011, a *pepO* knockout plasmid, was introduced  
418 in MG1363 via electroporation. Knockout mutants were obtained by a two-step homologous

419 recombination strategy (17), First, plasmid chromosomal integrates were selected on  
420 erythromycin containing GM17 plates. Subsequently, the marker-free knockout strain was  
421 obtained through counterselection on 5-FOA in SA medium plates. The resulting strain,  
422 CH011 (MG $\Delta$ pepO), underwent the same 2-step recombination protocol using pCH001 to  
423 obtain the peptidase double mutant strain CH017 (MG $\Delta$ pepOF<sub>2</sub>). Strain CH018  
424 (MG $\Delta$ pepOF2O2) was obtained using the strategy with plasmid pCH002 on strain CH017. All  
425 relevant chromosomal regions of each deletion strain were confirmed by nucleotide  
426 sequencing.

#### 427 **Construction of plasmid pCH020 for expressing proteinase PrtP in *L. lactis* MG1363**

428 The flanking regions of the *prtPM* genes from plasmid pLP712 (47) were amplified together  
429 using primers pCH-0173/pCH-0174. The fragment was ligated into plasmid pTLR employing  
430 NcoI/XhoI restriction sites. The resulting plasmid was named pCH020.

#### 431 **$\beta$ -casein degradation *in vivo***

432 *In vivo*  $\beta$ -casein breakdown was examined using the method of Kunji et al. (49) with the  
433 following modifications: an overnight culture was diluted to a starting optical density at 600  
434 nm (OD<sub>600</sub>) of 0.05 in 50 ml of GM17 with 5  $\mu$ g/ml erythromycin, when required. The culture  
435 was grown at 30 °C and when the OD<sub>600</sub> reached 0.7, which corresponds to the early  
436 exponential growth phase, the cells were collected by centrifugation at 6000 g for 5 min. They  
437 were washed twice with wash buffer (100 mM MES-KOH (pH6.5) with 2mM CaCl<sub>2</sub>) to prevent  
438 autolysis and release of the proteinase PrtP. Cells were then concentrated to an OD<sub>600</sub>  
439 of 14 and resuspended in 2 ml of 4 mg/ml  $\beta$ -casein (Cat # C6905, Sigma-Aldrich, Santa Clara,



440 CA, USA) in wash buffer with 0.5% (w/v) glucose. The suspension was incubated for 3 h at  
441 30 °C with slow rotation (10 rpm) in a rotator incubator oven (Cat # G2545A, Agilent  
442 Technologies, Inc., CA, USA). Cells were then spun down at 12000 g for 3 min after which both  
443 the supernatant and the cells were saved at -80 °C until further use.

#### 444 **Cytoplasmic peptidome extract preparation**

445 Frozen cells incubated with  $\beta$ -casein were thawed and resuspended in 2 ml 1 M LiCl in 50 mM  
446 Tris (pH 8.0) to extract proteins in the surface layer or anchored to the cell wall through non-  
447 covalent interactions (50). After incubation at 4 °C for 1 h, the cells were spun down at 4000 g  
448 for 10 min. The supernatant, named LiCl extract and saved at -80 °C. The cells were washed  
449 twice with MiliQ water (Milipore, MA, USA), and subsequently disrupted in a Mini-Beadbeater  
450 (Cat # 112011EUR, BioSpec, OK, USA) using 3 cycles of 1 min ON, 1 min OFF. Disrupted cells  
451 were spun down at 11000 xg at 4 °C for 10 min. The supernatant was collected and filtered  
452 through a 0.2  $\mu$ m pore-size filter (Cat # 41055511, Boom BV, Meppel, the Netherlands), and  
453 then ultrafiltrated through Amicon 3-KDa molecular weight cut-off membranes (Cat #  
454 UFC500324, Millipore, MA, USA). The pool of peptides less than 3-KDa was collected and  
455 saved at -80 °C until further use.

#### 456 **nanoLC-MS/MS**

457 All samples were analyzed at the Interfaculty Mass Spectrometry Center, University of  
458 Groningen, on a nanoLC-MS/MS consisting of an Ultimate 3000 LC system (Dionex,  
459 Amsterdam, the Netherlands) interfaced with a Q-Exactive plus mass spectrometer (Thermo  
460 Fisher Scientific, MA, USA). Peptide mixtures were loaded onto a 5 mm  $\times$  300  $\mu$ m i.d. C18

461 PepMAP100 trapping column (Thermo Fisher Scientific, MA, USA) with 2% acetonitrile in 0.1%  
462 formic acid at 20  $\mu$ L/min. After loading and washing for 3 min, peptides were eluted onto a 15  
463 cm  $\times$  75  $\mu$ m i.d. C18 PepMAP100 nanocolumn (Dionex, Amsterdam, the Netherlands). A  
464 mobile phase gradient at a flow rate of 300 nL/min and with a total run time of 75 min was  
465 used: 2% – 50% of solvent B in 60 min; 50% – 90% B in 1 min; 90% B during 13 min, and back  
466 to 2% B in 1 min (held for 15 min). Solvent A was 100:0 water/acetonitrile (v/v) with 0.1%  
467 formic acid, and solvent B was 0:100 water/acetonitrile (v/v) with 0.1% formic acid. In the  
468 nanospray source a stainless-steel emitter (Thermo Fisher Scientific, MA, USA) was used at a  
469 spray voltage of 1.8 kV with no sheath or auxiliary gas flow. The ion transfer tube temperature  
470 was 275  $^{\circ}$ C. Spectra were acquired in data-dependent mode with a survey scan at m/z 300 –  
471 1650 at a resolution of 70 000 followed by MS/MS fragmentation of the top 10 precursor ions.  
472 Singly charged ions were excluded from MS/MS experiments and fragmented precursor ions  
473 were dynamically excluded for 20 s. PEAKS Studio version X software (Bioinformatics  
474 Solutions, Inc., Waterloo, Canada) was used to search the MS data against a protein sequence  
475 database of the *L. lactis* MG1363 proteome (UniProt database) to which the sequence of  $\beta$ -  
476 casein (Uniprot P02666) was added. Search parameters: no enzyme specificity; fixed  
477 modification: carbamidomethylation of cysteine; variable modifications: oxidation of  
478 methionine and phosphorylation of serine; precursor mass tolerance of 15 ppm; fragment  
479 mass tolerance of 0.02 Da. The false discovery rate was set at 0.1%.

#### 480 **Data analysis and visualization**

481 The peptide spectrum matches (PSM), identified peptides, and identified protein number  
482 were obtained from PEAKS. The identified peptides were exported for further analysis in R by

483 using Pheatmap R package. Venn diagrams were made using the webtool Calculate and Draw  
484 Custom Venn Diagrams (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Gene ontology  
485 analysis was performed using Gene Set Enrichment Analysis (GSEA) provided by the  
486 GENOME2D software available at <http://server.molgenrug.nl/index.php/gsea-pro>. Circos plots  
487 were made using the Circlize R package (51). For further peptidomic analysis, peptide  
488 sequences of biological triplicate were combined, and only unique peptides present at least  
489 twice among triplicates were used. Thus, 7 datasets of all 21 samples were generated. The  
490 further data analysis mentioned below was performed on these combined datasets. Profiles  
491 of  $\beta$ -casein-derived peptides in the *L. lactis* intracellular peptidome were visualized by the  
492 web-based tool Peptigram (30). Data of the parameter relative intensity were generated from  
493 the average intensity of the replicates of each chosen peptide. Peptide physicochemical  
494 properties were computed using the aminoAcidProperties function of the R package  
495 “alakazam”, version 0.2.8 (52). The proportion of proline in each peptide was manually  
496 calculated. The Upset plot of our optimized bioactive peptide database was generated using  
497 the Upset R package (53).  
498

499 **References**

- 500 1. Sauer M, Russmayer H, Grabherr R, Peterbauer CK, Marx H. 2017. The Efficient Clade:  
501 Lactic Acid Bacteria for Industrial Chemical Production. *Trends Biotechnol* 35:756–769.
- 502 2. Mierau I, Venema G, Kok J, Kunji ERS. 1997. Casein and peptide degradation in lactic  
503 acid bacteria. *Biotechnol Genet Eng Rev* 14:279–302.
- 504 3. Venegas-Ortega MG, Flores-Gallegos AC, Martínez-Hernández JL, Aguilar CN, Nevárez-  
505 Moorillón G V. 2019. Production of Bioactive Peptides from Lactic Acid Bacteria: A  
506 Sustainable Approach for Healthier Foods. *Compr Rev Food Sci Food Saf* 18:1039–1051.
- 507 4. Guédon E, Renault P, Ehrlich SD, Delorme C. 2001. Transcriptional pattern of genes  
508 coding for the proteolytic system of *Lactococcus lactis* and evidence for coordinated  
509 regulation of key enzymes by peptide supply. *J Bacteriol* 183:3614–3622.
- 510 5. Den Hengst CD, Van Hijum SAFT, Geurts JMW, Nauta A, Kok J, Kuipers OP. 2005. The  
511 *Lactococcus lactis* CodY regulon: Identification of a conserved cis-regulatory element. *J*  
512 *Biol Chem* 280:34332–34342.
- 513 6. Hernández-Ledesma B, Amigo L, Ramos M, Recio I. 2004. Release of angiotensin  
514 converting enzyme-inhibitory peptides by simulated gastrointestinal digestion of  
515 infant formulas. *Int Dairy J* 14:889–898.
- 516 7. Shanmugam VP, Kapila S, Sonfack TK, Kapila R. 2015. Antioxidative peptide derived  
517 from enzymatic digestion of buffalo casein. *Int Dairy J* 42:1–5.

- 518 8. Otte J, Shalaby SMA, Zakora M, Nielsen MS. 2007. Fractionation and identification of  
519 ACE-inhibitory peptides from  $\alpha$ -lactalbumin and  $\beta$ -casein produced by thermolysin-  
520 catalysed hydrolysis. *Int Dairy J* 17:1460–1472.
- 521 9. Hayes M, Stanton C, Slattery H, O’Sullivan O, Hill C, Fitzgerald GF, Ross RP. 2007. Casein  
522 fermentate of *Lactobacillus animalis* DPC6134 contains a range of novel propeptide  
523 angiotensin-converting enzyme inhibitors. *Appl Environ Microbiol* 73:4658–4667.
- 524 10. Gobbetti M, Ferranti P, Smacchi E, Goffredi F, Addeo F. 2000. Production of angiotensin-  
525 I-converting-enzyme-inhibitory peptides in fermented milks started by *Lactobacillus*  
526 *delbrueckii* subsp. *bulgaricus* SS1 and *Lactococcus lactis* subsp. *cremoris* FT4. *Appl*  
527 *Environ Microbiol* 66:3898–3904.
- 528 11. Quirós A, Ramos M, Muguerza B, Delgado MA, Miguel M, Aleixandre A, Recio I. 2007.  
529 Identification of novel antihypertensive peptides in milk fermented with *Enterococcus*  
530 *faecalis*. *Int Dairy J* 17:33–41.
- 531 12. Toldrá F, Nollet LML. 2013. Proteomics in foods: Principles and applications Proteomics  
532 in Foods: Principles and Applications.
- 533 13. Patejko M, Jacyna J, Markuszewski MJ. 2017. Sample preparation procedures utilized in  
534 microbial metabolomics: An overview. *J Chromatogr B Anal Technol Biomed Life Sci*  
535 1043:150–157.
- 536 14. Schrader M, Fricker L. 2018. Peptidomics. Springer.

- 537 15. Wegmann U, Overweg K, Jeanson S, Gasson M, Shearman C. 2012. Molecular  
538 characterization and structural instability of the industrially important composite  
539 metabolic plasmid pLP712. *Microbiology* 158:2936–2945.
- 540 16. Haandrikman AJ, Meesters R, Laan H, Konings WN, Kok J, Venema G. 1991. Processing of  
541 the lactococcal extracellular serine proteinase. *Appl Environ Microbiol* 57:1899–1904.
- 542 17. Solem C, Defoor E, Jensen PR, Martinussen J. 2008. Plasmid pCS1966, a new  
543 selection/counterselection tool for lactic acid bacterium strain construction based on  
544 the *oroP* gene, encoding an orotate transporter from *Lactococcus lactis*. *Appl Environ*  
545 *Microbiol* 74:4772–4775.
- 546 18. van der Meulen SB. 2018. RNA regulation in *Lactococcus lactis*. Rijksuniversiteit  
547 Groningen.
- 548 19. Pinu FR, Villas-Boas SG, Aggio R. 2017. Analysis of intracellular metabolites from  
549 microorganisms: Quenching and extraction protocols. *Metabolites* 7.
- 550 20. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, Zhang W, Zhang Z, Lajoie GA, Ma B. 2012.  
551 PEAKS DB: *De Novo* Sequencing Assisted Database Search for Sensitive and Accurate  
552 Peptide Identification. *Mol Cell Proteomics* 11:M111.010587.
- 553 21. Vaudel M, Barsnes H, Berven FS, Sickmann A, Martens L. 2011. SearchGUI: An open-  
554 source graphical user interface for simultaneous OMSSA and X! Tandem searches.  
555 *Proteomics* 11:996–999.
- 556 22. Park CY, Klammer AA, Käll L, MacCoss MJ, Noble WS. 2008. Rapid and accurate peptide  
557 identification from tandem mass spectra. *J Proteome Res* 7:3022–3027.

- 558 23. Eng JK, Jahan TA, Hoopmann MR. 2013. Comet: an open-source MS/MS sequence  
559 database search tool. *Proteomics* 13:22–24.
- 560 24. Fenyö D, Beavis RC. 2003. A method for assessing the statistical significance of mass  
561 spectrometry-based protein identifications using general scoring schemes. *Anal Chem*  
562 75:768–774.
- 563 25. Tabb DL, Fernando CG, Chambers MC. 2007. MyriMatch: highly accurate tandem mass  
564 spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res*  
565 6:654–661.
- 566 26. Kim S, Pevzner PA. 2014. MS-GF+ makes progress towards a universal database search  
567 tool for proteomics. *Nat Commun* 5:5277.
- 568 27. Dorfer V, Pichler P, Stranzl T, Stadlmann J, Taus T, Winkler S, Mechtler K. 2014. MS  
569 Amanda, a universal identification algorithm optimized for high accuracy tandem mass  
570 spectra. *J Proteome Res* 13:3679–3684.
- 571 28. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH.  
572 2004. Open mass spectrometry search algorithm. *J Proteome Res* 3:958–964.
- 573 29. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V, Mann M. 2011. Andromeda: a  
574 peptide search engine integrated into the MaxQuant environment. *J Proteome Res*  
575 10:1794–1805.
- 576 30. Manguy J, Jehl P, Dillon ET, Davey NE, Shields DC, Holton TA. 2017. Peptigram: a web-  
577 based application for peptidomics data visualization. *J Proteome Res* 16:712–719.

- 578 31. Huang C, Hernandez-Valdes JA, Kuipers OP, Kok J. 2020. Lysis of an *Lactococcus lactis*  
579 Di-peptidase Mutant and Rescue by Mutation in the Pleiotropic Regulator CodY. *Appl*  
580 *Environ Microbiol.*
- 581 32. Minkiewicz P, Dziuba J, Iwaniak A, Dziuba M, Darewicz M. 2008. BIOPEP database and  
582 other programs for processing bioactive peptide sequences. *J AOAC Int* 91:965–980.
- 583 33. Zamyatnin AA, Borchikov AS, Vladimirov MG, Voronina OL. 2006. The EROP-Moscow  
584 oligopeptide database. *Nucleic Acids Res* 34:D261–D266.
- 585 34. Nielsen SD, Beverly RL, Qu Y, Dallas DC. 2017. Milk bioactive peptide database: A  
586 comprehensive database of milk protein-derived bioactive peptides and novel  
587 visualization. *Food Chem* 232:673–682.
- 588 35. Yoshikawa M, Tani F, Yoshimura T, Chiba H. 1986. Opioid peptides from milk proteins.  
589 *Agric Biol Chem* 50:2419–2421.
- 590 36. Osborne S, Chen W, Addepalli R, Colgrave M, Singh T, Tran C, Day L. 2014. In vitro  
591 transport and satiety of a beta-lactoglobulin dipeptide and beta-casomorphin-7 and its  
592 metabolites. *Food Funct* 5:2706–2718.
- 593 37. Dubynin VA, Asmakova LS, Sokhanenkova NY, Bessalova ZD, Nezavibat'ko VN,  
594 Kamenskii AA. 1998. Comparative analysis of neurotropic activity of exorphines,  
595 derivatives of dietary proteins. *Bull Exp Biol Med* 125:131–134.
- 596 38. Hatzoglou A, Bakogeorgou E, Hatzoglou C, Martin PM, Castanas E. 1996.  
597 Antiproliferative and receptor binding properties of  $\alpha$ - and  $\beta$ -casomorphins in the T47D  
598 human breast cancer cell line. *Eur J Pharmacol* 310:217–223.



- 599 39. Villegas JM, Picariello G, Mamone G, Espeche Turbay MB, Savoy de Giori G, Hebert EM.  
600 2014. Milk-derived angiotensin-I-converting enzymeinhibitory peptides generated by  
601 *Lactobacillus delbrueckii* subsp. *lactis* CRL 581. *Peptidomics* 1:22–29.
- 602 40. Yamamoto N, Akino A, Takano T. 1994. Antihypertensive Effect of the Peptides Derived  
603 from Casein by an Extracellular Proteinase from *Lactobacillus helveticus* CP790. *J Dairy*  
604 *Sci* 77:917–922.
- 605 41. Shinoda I, Tada M, Fukui S. 1986. Bitter taste of h-pro-phe-pro-gly-pro-ile-pro-oh  
606 corresponding to the partial sequence (Positions 61-67) of bovine  $\beta$ -casein, and related  
607 peptidest. *Agric Biol Chem* 50:1247–1254.
- 608 42. Xu Q, Hong H, Wu J, Yan X. 2019. Bioavailability of bioactive peptides derived from food  
609 proteins across the intestinal epithelial membrane: A review. *Trends Food Sci Technol*  
610 86:399–411.
- 611 43. Detmers FJM, Lanfermeijer FC, Abele R, Jack RW, Tampé R, Konings WN, Poolman B.  
612 2000. Combinatorial peptide libraries reveal the ligand-binding mechanism of the  
613 oligopeptide receptor OppA of *Lactococcus lactis*. *Proc Natl Acad Sci U S A* 97:12487–  
614 12492.
- 615 44. Monnet V, Nardi M, Chopin A, Chopin MC, Gripon JC. 1994. Biochemical and Genetic-  
616 Characterization of Pepf, an Oligopeptidase from *Lactococcus-Lactis*. *J Biol Chem*  
617 269:32070–32076.

- 618 45. Van Die IM, Bergmans HEN, Hoekstra WPM. 1983. Transformation in *Escherichia coli*:  
619 studies on the role of the heat shock in induction of competence. *Microbiology*  
620 129:663–670.
- 621 46. Holo H, Nes IF. 1989. High-frequency transformation, by electroporation, of  
622 *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media.  
623 *Appl Environ Microbiol* 55:3119–3123.
- 624 47. Wegmann U, Overweg K, Jeanson S, Gasson M, Shearman C. 2012. Molecular  
625 characterization and structural instability of the industrially important composite  
626 metabolic plasmid pLP712. *Microbiol (United Kingdom)* 158:2936–2945.
- 627 48. Wegmann U, O'Connell-Motherway M, Zomer A, Buist G, Shearman C, Canchaya C,  
628 Ventura M, Goesmann A, Gasson MJ, Kuipers OP, van Sinderen D, Kok J. 2007. Complete  
629 Genome Sequence of the Prototype Lactic Acid Bacterium *Lactococcus lactis* subsp.  
630 *cremoris* MG1363. *J Bacteriol* 189:3256–3270.
- 631 49. Kunji ERS, Fang G, Jeronimus-Stratingh CM, Bruins AP, Poolman B, Konings WN. 1998.  
632 Reconstruction of the proteolytic pathway for use of  $\beta$ -casein by *Lactococcus lactis*. *Mol*  
633 *Microbiol* 27:1107–1118.
- 634 50. Koval SF, Murray RG. 1984. The isolation of surface array proteins from bacteria. *Can J*  
635 *Biochem Cell Biol* 62:1181–1189.
- 636 51. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize implements and enhances circular  
637 visualization in R. *Bioinformatics* 30:2811–2812.

- 638 52. Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. 2015.  
639 Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire  
640 sequencing data. *Bioinformatics* 31:3356–3358.
- 641 53. Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of  
642 intersecting sets and their properties. *Bioinformatics* 33:2938–2940.
- 643 54. Gasson MJ. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other  
644 lactic streptococci after protoplast-induced curing. *J Bacteriol* 154:1–9.
- 645 55. Mierau I, Kunji ER, Leenhouts KJ, Hellendoorn MA, Haandrikman AJ, Poolman B, Konings  
646 WN, Venema G, Kok J. 1996. Multiple-peptidase mutants of *Lactococcus lactis* are  
647 severely impaired in their ability to grow in milk. *J Bacteriol* 178:2794–2803.
- 648 56. Taylor RG, Walker DC, McInnes RR. 1993. *E.coli* host strains. *Nucleic Acid Res* 21:1677–  
649 1678.
- 650 57. Saito T, Nakamura T, Kitazawa H, Kawai Y, Itoh T. 2000. Isolation and structural analysis  
651 of antihypertensive peptides that exist naturally in Gouda cheese. *J Dairy Sci* 83:1434–  
652 1440.
- 653 58. Uenishi H, Kabuki T, Seto Y, Serizawa A, Nakajima H. 2012. Isolation and identification of  
654 casein-derived dipeptidyl-peptidase 4 (DPP-4)-inhibitory peptide LPQNIPPL from  
655 gouda-type cheese and its effect on plasma glucose in rats. *Int Dairy J* 22:24–30.
- 656 59. Jinsmaa Y, Yoshikawa M. 1999. Enzymatic release of neocasomorphin and  $\beta$ -  
657 casomorphin from bovine  $\beta$ -casein. *Peptides* 20:957–962.

- 658 60. Sabeena Farvin KH, Baron CP, Nielsen NS, Otte J, Jacobsen C. 2010. Antioxidant activity  
659 of yoghurt peptides: Part 2 - Characterisation of peptide fractions. *Food Chem*  
660 123:1090–1097.
- 661 61. Sedaghati M, Ezzatpanah H, Mashhadi Akbar Boojar M, Tajabadi Ebrahimi M, Kobarfard  
662 F. 2016. Isolation and identification of some antibacterial peptides in the plasmin-  
663 digest of  $\beta$ -casein. *LWT - Food Sci Technol* 68:217–225.
- 664 62. Perpetuo EA, Juliano L, Lebrun I. 2003. Biochemical and Pharmacological Aspects of  
665 Two Bradykinin-Potentiating Peptides Obtained from Tryptic Hydrolysis of Casein. *J*  
666 *Protein Chem* 22:601–606.
- 667 63. Lu Y, Govindasamy-Lucey S, Lucey JA. 2016. Angiotensin-I-converting enzyme-  
668 inhibitory peptides in commercial Wisconsin Cheddar cheeses of different ages. *J Dairy*  
669 *Sci* 99:41–52.
- 670 64. Eisele T, Stressler T, Kranz B, Fischer L. 2013. Bioactive peptides generated in an enzyme  
671 membrane reactor using *Bacillus lentus* alkaline peptidase. *Eur Food Res Technol*  
672 236:483–490.
- 673 65. Smacchi E, Gobetti M. 1998. Peptides from several Italian cheeses inhibitory to  
674 proteolytic enzymes of lactic acid bacteria, *pseudomonas fluorescens* ATCC 948 and to  
675 the angiotensin I-converting enzyme. *Enzyme Microb Technol* 22:687–694.
- 676 66. Rival SG, Boeriu CG, Wichers HJ. 2001. Caseins and casein hydrolysates. 2. Antioxidative  
677 properties and relevance to lipoxygenase inhibition. *J Agric Food Chem* 49:295–302.

- 678 67. Robert MC, Razaname A, Mutter M, Juillerat MA. 2004. Identification of angiotensin-I-  
679 converting enzyme inhibitory peptides derived from sodium caseinate hydrolysates  
680 produced by *Lactobacillus helveticus* NCC 2765. *J Agric Food Chem* 52:6923–6931.
- 681 68. Maeno M, Yamamoto N, Takano T. 1996. Identification of an antihypertensive peptide  
682 from casein hydrolysate produced by a proteinase from *Lactobacillus helveticus* CP790.  
683 *J Dairy Sci* 79:1316–1321.
- 684 69. Birkemo GA, O’Sullivan O, Ross RP, Hill C. 2009. Antimicrobial activity of two peptides  
685 casecidin 15 and 17, found naturally in bovine colostrum. *J Appl Microbiol* 106:233–240.
- 686 70. Rojas-Ronquillo R, Cruz-Guerrero A, Flores-Nájera A, Rodríguez-Serrano G, Gómez-Ruiz  
687 L, Reyes-Grajeda JP, Jiménez-Guzmán J, García-Garibay M. 2012. Antithrombotic and  
688 angiotensin-converting enzyme inhibitory properties of peptides released from bovine  
689 casein by *Lactobacillus casei* Shirota. *Int Dairy J* 26:147–154.
- 690 71. Sandré C, Gleizes A, Forestier F, Gorges-Kergot R, Chilmonczyk S, Léonil J, Moreau M-C,  
691 Labarre C. 2001. A Peptide Derived from Bovine  $\beta$ -Casein Modulates Functional  
692 Properties of Bone Marrow-Derived Macrophages from Germfree and Human Flora-  
693 Associated Mice. *J Nutr* 131:2936–2942.
- 694 72. Maruyama S, Nakagomi K, Tomizuka N, Suzuki H. 1985. Angiotensin i-converting  
695 enzyme inhibitor derived from an enzymatic hydrolysate of casein. II. isolation and  
696 bradykinin-potentiating activity on the uterus and the ileum of rats. *Agric Biol Chem*  
697 49:1405–1409.
- 698



700 **Table 1. Strains used in the peptidomic experiment.**

Strain	Species	Description	Short name	Ab <sup>r</sup>	Reference
MG1363	<i>L. lactis</i>	Plasmid free derivative of NCD0712	MG1363		(54)
CH000	<i>L. lactis</i>	MG1363 with plasmid pCH020	MG1363(PrtP <sup>r</sup> )	Ery	this study
IM14	<i>L. lactis</i>	MG1363Δ <i>pepNXOTC</i>	MGΔ <i>pepNXOTC</i>		(55)
CH001	<i>L. lactis</i>	MG1363Δ <i>pepF</i> <sub>2</sub>	MGΔ <i>pepF</i> <sub>2</sub>		this study
CH002	<i>L. lactis</i>	MG1363Δ <i>pepO</i> <sub>2</sub>	MGΔ <i>pepO</i> <sub>2</sub>		this study
CH003	<i>L. lactis</i>	MG1363Δ <i>pepA</i>	MGΔ <i>pepA</i>		this study
CH004	<i>L. lactis</i>	MG1363Δ <i>pepP</i>	MGΔ <i>pepP</i>		this study
CH005	<i>L. lactis</i>	MG1363Δ <i>pepV</i>	MGΔ <i>pepV</i>		this study
CH007	<i>L. lactis</i>	MG1363Δ <i>pcp</i>	MGΔ <i>pcp</i>		this study
CH008	<i>L. lactis</i>	MG1363Δ <i>pepQ</i>	MGΔ <i>pepQ</i>		this study
CH009	<i>L. lactis</i>	MG1363Δ <i>pepD</i> <sub>A</sub>	MGΔ <i>pepD</i> <sub>A</sub>		this study
CH010	<i>L. lactis</i>	MG1363 (pLP712Δ <i>pepF</i> <sub>1</sub> )	MGΔ <i>pepF</i> <sub>1</sub>		this study
CH011	<i>L. lactis</i>	MG1363Δ <i>pepO</i>	MGΔ <i>pepO</i>		this study
CH012	<i>L. lactis</i>	MG1363Δ <i>pepC</i>	MGΔ <i>pepC</i>		this study
CH013	<i>L. lactis</i>	MG1363Δ <i>pepN</i>	MGΔ <i>pepN</i>		this study
CH014	<i>L. lactis</i>	MG1363Δ <i>pepX</i>	MGΔ <i>pepX</i>		this study
CH015	<i>L. lactis</i>	MG1363Δ <i>pepT</i>	MGΔ <i>pepT</i>		this study
CH016	<i>L. lactis</i>	MG1363Δ <i>pepD</i> <sub>B</sub>	MGΔ <i>pepD</i> <sub>B</sub>		this study
CH017	<i>L. lactis</i>	MG1363Δ <i>pepO</i> <sub>F</sub> <sub>2</sub>	MGΔ <i>pepO</i> <sub>F</sub> <sub>2</sub>		this study
CH018	<i>L. lactis</i>	MG1363Δ <i>pepO</i> <sub>F</sub> <sub>2</sub> <sub>O</sub> <sub>2</sub>	MGΔ <i>pepO</i> <sub>F</sub> <sub>2</sub> <sub>O</sub> <sub>2</sub>		this study
CH019	<i>L. lactis</i>	MG1363Δ <i>pepV</i> <sub>D</sub> <sub>A</sub>	MGΔ <i>pepV</i> <sub>D</sub> <sub>A</sub>		this study
CH020	<i>L. lactis</i>	MG1363Δ <i>pepV</i> <sub>D</sub> <sub>A</sub> <sub>T</sub>	MGΔ <i>pepV</i> <sub>D</sub> <sub>A</sub> <sub>T</sub>		this study
CH021	<i>L. lactis</i>	MG1363Δ <i>pepV</i> <sub>D</sub> <sub>A</sub> <sub>TD</sub> <sub>B</sub>	MGΔ <i>pepV</i> <sub>D</sub> <sub>A</sub> <sub>TD</sub> <sub>B</sub>		this study
CH022	<i>L. lactis</i>	MG1363Δ <i>pepAN</i>	MGΔ <i>pepAN</i>		this study
CH023	<i>L. lactis</i>	MG1363Δ <i>pepANC</i>	MGΔ <i>pepANC</i>		this study
CH024	<i>L. lactis</i>	MG1363Δ <i>pepANC</i> <sub>pcp</sub>	MGΔ <i>pepANC</i> <sub>pcp</sub>		this study
CH025	<i>L. lactis</i>	MG1363Δ <i>pepXP</i>	MGΔ <i>pepXP</i>		this study
CH026	<i>L. lactis</i>	MG1363Δ <i>pepXPQ</i>	MGΔ <i>pepXPQ</i>		this study
CH027	<i>L. lactis</i>	MG1363Δ <i>pepNXOTCF</i> <sub>2</sub>	MGΔ <i>pepNXOTCF</i> <sub>2</sub>		this study
CH028	<i>L. lactis</i>	MG1363Δ <i>pepNXOTCF</i> <sub>2</sub> <sub>O</sub> <sub>2</sub>	MGΔ <i>pepNXOTCF</i> <sub>2</sub> <sub>O</sub> <sub>2</sub>		this study
CH029	<i>L. lactis</i>	MG1363Δ <i>pepNXOTCV</i>	MGΔ <i>pepNXOTCV</i>		this study
CH030	<i>L. lactis</i>	MG1363Δ <i>pepNXOTCVD</i> <sub>A</sub>	MGΔ <i>pepNXOTCVD</i> <sub>A</sub>		this study
CH031	<i>L. lactis</i>	CH020 with plasmid pCH018	MGΔ <i>pepO</i> <sub>F</sub> <sub>2</sub> <sub>O</sub> <sub>2</sub> (PrtP <sup>r</sup> )	Ery	this study
CH032	<i>L. lactis</i>	CH021 with plasmid pCH020	MGΔ <i>pepV</i> <sub>D</sub> <sub>A</sub> <sub>TD</sub> <sub>B</sub> (PrtP <sup>r</sup> )	Ery	this study

Strain	Species	Description	Short name	Ab <sup>r</sup>	Reference
CH033	<i>L. lactis</i>	CH024 with plasmid pCH020	MGΔ <i>pepANCpcp</i> (PrtP <sup>+</sup> )	Ery	this study
CH034	<i>L. lactis</i>	CH026 with plasmid pCH020	MGΔ <i>pepXPQ</i> (PrtP <sup>+</sup> )	Ery	this study
CH035	<i>L. lactis</i>	CH028 with plasmid pCH020	MGΔ <i>pepNXOTCF<sub>2</sub>O<sub>2</sub></i> (PrtP <sup>+</sup> )	Ery	this study
CH036	<i>L. lactis</i>	CH030 with plasmid pCH020	MGΔ <i>pepNXOTCVD<sub>Δ</sub></i> (PrtP <sup>+</sup> )	Ery	this study
DH5α	<i>E. coli</i>	<i>fhuA2 lac(del)U169 phoA glnV44</i> <i>Φ80' lacZ(del)M15 gyrA96 recA1</i> <i>relA1 endA1 thi-1 hsdR17</i>			(56)

701 Ab<sup>r</sup>=Antibiotic resistance. Ery = erythromycin

702



703 **Table 2. Bovine  $\beta$ -casein-derived bioactive peptides identified in the intracellular**  
 704 **peptidome of *L. lactis* MG1363 and its peptidase knockout mutants.**

Peptide Sequence	$\beta$ -CN fragment (start-end)	Theoretical Mass (Da)	Mass Error (ppm)	Bioactivity in Reference	Ref NO.
<b>MG1363</b>					
YFPFGPIPN	75-83	1000.5018	0.9 / 3.4 / -1.8	ACE-I; DPP-IV-I; Opioid	(57–59)
LPQNIPP	85-91	777.4385	3.3 / 4.9 / 3.6	DPP-IV-I	(58)
LPQNIPPL	85-92	890.5225	3.1 / - / -0.5	DPP-IV-I	(58)
TQTPVVVPFLQPE	93-106	1550.8344	3.6 / 6.8 / -	Anti-O	(60)
<b>MG<math>\Delta</math>pepOF<sub>2</sub>O<sub>2</sub></b>					
YFPFGPIPN	75-83	1000.5018	3.3 / 2.8 / 0	ACE-I; DPP-IV-I; Opioid	(57–59)
HKEMPFPK	121-128	1012.5164	0.3 / -2 / -1.3	Anti-M	(61)
YPVEPFTE	129-136	980.4491	9.7 / 8 / 4.3	ACE-I	(62)
SQSKVLPVPQ	181-190	1081.6132	6.7 / 6.9 / 5.4	ACE-I	(9)
RDMPIQAF	198-205	976.4800	- / 6.1 / 5	ACE-I	(40)
QEPVLGPVRGPFPIIV	209-224	1716.9926	4.8 / 6.1 / 2.5	ACE-I	(63)
<b>MG<math>\Delta</math>pepNXOTCF<sub>2</sub>O<sub>2</sub></b>					
LNVGGEIVE	21-29	968.5178	2.5 / 3.8 / 2.5	ACE-I	(10)
VYFPFGPIPN	74-83	1099.5702	1.3 / 2 / 3.2	ACE-I; Anti-O	(64)
LVYFPFGPIPNLSLPQ	73-87	1637.8817	3.6 / 3.8 / -	ACE-I; PEP-I	(65)
LPQNIPPL	85-92	890.5225	2.8 / 1.5 / 4	DPP-IV-I	(58)
PQNIPPL	86-92	777.4385	2.9 / 2 / 2.2	DPP-IV-I	(58)
NIPPLTQTPV	88-97	1078.6023	4.2 / 4.6 / 4/6	ACE-I	(10)
TQTPVVVPFLQPE	93-106	1550.8344	2.5 / 2.5 / 2.5	Anti-O	(60)
VKEAMAPK	113-120	872.4789	-2.9 / -4.1 / -1.6	Anti-O; Anti-M	(66)
HKEMPFPK	121-128	1012.5164	5.2 / 2.9 / 4.1	Anti-M	(61)
LHLPLPL	148-154	801.5112	2.8 / 5.5 / 6.1	ACE-I	(11)
NLHLPLPLL	147-155	1028.6382	1.9 / 3.8 / 3.9	ACE-I	(67)
SQSKVLPVPQ	181-190	1081.6132	3.8 / 5.1 / 4.5	ACE-I	(9)
KVLPVPQK	184-191	907.5854	5.1 / 4.5 / 7	Anti-O	(66)
KVLPVPQ	184-190	779.4905	3.6 / 1.8 / 9.1	ACE-I	(68)
VLPVPQK	185-191	779.4905	4.7 / 6.9 / 7.6	Anti-M; Anti-O	(66)
RDMPIQAF	198-205	976.4800	4.7 / 3.8 / 4.9	ACE-I	(40)
YQEPVLGPVRGPFPIIV	208-224	1880.0559	4.3 / 4.3 / -	ACE-I; Anti-M; Immuno-R	(69–71)
QEPVLGPVRGPFPIIV	209-224	1716.9926	2.6 / 2.7 / 2.2	ACE-I	(63)
<b>MG<math>\Delta</math>pepXPQ</b>					
YFPFGPIPN	85-83	1000.5018	-0.6 / -1.4 / -0.3	ACE-I; DPP-IV-I; Opioid	(57–59)
LPQNIPPL	85-92	890.5225	1.5 / 2.8 / -	DPP-IV-I	(58)

Peptide Sequence	$\beta$ -CN fragment (start-end)	Theoretical Mass (Da)	Mass Error (ppm)	Bioactivity in Reference	RefNO.
<b>MG<math>\Delta</math>pepVD<sub>A</sub>TD<sub>B</sub></b>					
-					
<b>MG<math>\Delta</math>pepNXOTCVD<sub>A</sub></b>					
LNVPGEIVE	21-29	968.5178	- / 1.9 / 2	ACE-I	(10)
VYPFGPIP	74-83	1099.5702	- / 2.5 / 3.4	ACE-I; Anti-O	(64)
LPQNIPPL	85-92	890.5225	- / 3.9 / 4.9	DPP-IV-I	(58)
PQNIPPL	86-92	777.4385	- / 1.7 / 4.1	DPP-IV-I	(58)
NIPPLTQTPV	88-97	1078.6023	- / 3.9 / 5.5	ACE-I	(10)
TQTPVVPPFLQPE	93-106	1550.8344	- / 3.6 / 4.7	Anti-O	(60)
LHLPLPL	148-154	801.5112	- / 5.8 / 3.5	ACE-I	(11)
SQSKVLPVPQ	181-190	1081.6132	- / 4.5 / 4.3	ACE-I	(9)
KVLPVPQ	184-190	779.4905	1.4 / 0.8 / 2.9	ACE-I	(68)
<b>MG<math>\Delta</math>pepANCpcp</b>					
LNVPGEIVE	21-29	968.5178	2.9 / 6.4 / 4.2	ACE-I	(10)
VYPFGPIP	74-83	1099.5702	5.4 / 5.5 / 4.8	ACE-I; Anti-O	(64)
NIPPLTQTPV	88-97	1078.6023	4.7 / 8 / 10	ACE-I	(10)
TQTPVVPPFLQPE	93-106	1550.8344	6.2 / 8.2 / 6.7	Anti-O	(60)
HKEMPPFK	121-128	1012.5164	0.1 / 1.2 / -0.1	Anti-M	(61)
LHLPLPL	148-154	801.5112	5.2 / 9.4 / 6.7	ACE-I	(11)
SQSKVLPVPQ	181-190	1081.6132	4.7 / 4.9 / 6.5	ACE-I	(9)
KVLPVPQ	184-190	779.4905	4.1 / 1.8 / 3.9	ACE-I	(68)
AVPYPQR	192-198	829.4446	2.9 / 1.4 / -	ACE-I; Anti-M; Anti-O	(61, 66, 72)

705  $\beta$ -CN,  $\beta$ -casein; ACE-I, angiotensin-converting-enzyme inhibitory; Anti-M, antimicrobial; Anti-  
706 O, antioxidative; DPP-IV-I, dipeptidyl peptidase 4 inhibitory; Immuno-R, immunoregulatory.  
707 The Mass Error is shown in parts per million (ppm) and is calculated as  $10^6 \times (\text{observed mass} -$   
708  $\text{theoretical mass}) / \text{theoretical mass}$ . The three numbers in the Mass Error column represent  
709 the biological triplicates; dash symbol (-) signifies that the peptide was not detected in one of  
710 the triplicates.

711

712 **Table 3. Plasmids used in this study.**

Plasmid	Host	Description	Ab <sup>r</sup>	reference
pCS1966	<i>E. coli</i>	<i>L. lactis</i> integration vector	Ery	(17)
pTLR	<i>E. coli</i>	<i>L. lactis</i> expression shuttle vector	Ery	Lab collection
pLP712	<i>L. lactis</i>	Prt <sup>+</sup> Lac <sup>+</sup> , 54kb proteinase/lactose plasmid of NCDO712		(54)
pCS1966	<i>E. coli</i>	<i>L. lactis</i> integration vector	Ery	this study
pCH001	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepF</i> <sub>2</sub>	Ery	this study
pCH002	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepO</i> <sub>2</sub>	Ery	this study
pCH003	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepA</i>	Ery	this study
pCH004	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepP</i>	Ery	this study
pCH005	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepV</i>	Ery	this study
pCH006	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepM</i>	Ery	this study
pCH007	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pcp</i>	Ery	this study
pCH008	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepQ</i>	Ery	this study
pCH009	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepD</i> <sub>A</sub>	Ery	this study
pCH010	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepF</i> <sub>1</sub> in pLP712	Ery	this study
pCH011	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepO</i>	Ery	this study
pCH012	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepC</i>	Ery	this study
pCH013	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepN</i>	Ery	this study
pCH014	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepX</i>	Ery	this study
pCH015	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepT</i>	Ery	this study
pCH016	<i>E. coli</i>	<i>L. lactis</i> integration vector, knockout <i>pepD</i> <sub>B</sub>	Ery	this study
pCH020	<i>E. coli</i> & <i>L. lactis</i>	pTLR-PrtPM, for expression of protease PrtP and PrtM under its own promoter	Ery	this study

713 Ab<sup>r</sup>=Antibiotic resistance. Ery = erythromycin.

714

715 **Figure 1 Peptidomics and bioactive peptide discovery in *L. lactis*.** (A) Schematic  
716 representation of the mutated proteolytic system of *L. lactis* MG1363 with the cell envelope-  
717 associated proteinase (PrtP).  $\beta$ -casein hydrolysis is initiated by PrtP, after its autoproteolytic  
718 activation with the aid of PrtM (16)). Subsequently, the oligopeptides are transported into the  
719 cells by the Opp transport system while di-/tripeptides are internalized by the Dpp or DtpT  
720 transport system. The peptides are then degraded by the concerted action of 15 peptidases,  
721 which are classified and colored by their indicated cleavage specificity. When a specific  
722 combination of peptidase genes is deleted, certain peptides will stay intact. Theoretically,  
723 small (bioactive) peptides, instead of free amino acids, will accumulate intracellularly. (B)  
724 General sketch of the peptidase gene knockout vector. The integration vector is based on  
725 plasmid pCS1966 (17). It contains an erythromycin-resistance gene (*ery*), orotate transporter  
726 gene (*oroP*), and a knockout cassette carrying two homology regions, one containing the  
727 region upstream of the gene to be deleted (HRU), the other encompassing a region  
728 downstream thereof (HRD), for double crossover integration at a certain peptidase gene locus.  
729

730 **Figure 2 *L. lactis* intracellular peptidomic sample preparation and data analysis**  
731 **optimization.** (A) Sample preparation workflow (for details, see M&M). Cells from  
732 exponentially growing *L. lactis* MG1363 and its peptidase knock-out mutants were incubated  
733 under slow rotation (180 rpm/min) in a  $\beta$ -casein solution (4 mg/ml) for 3.5 hours, after which  
734 they were disrupted by minibead-beating. The intracellular proteome was extracted and  
735 passed through a 3-kDa molecular weight cut-off filter. The filtrate, which is the intracellular  
736 peptidome, was desalted using C18 Zip-tips and analyzed by nanoLC-MS/MS. (B)  
737 Computational data analysis. Mass spectrum data (MS.raw) output and the proteome of *L.*  
738 *lactis* MG1363 plus  $\beta$ -casein sequence (both obtained from Uniprot) were used as inputs for  
739 PEAKS studio analysis. The output from PEAKS was further visualized through R programming  
740 and the peptidomics visualization web server Peptigram. (C) Comparison of 9 commonly used  
741 search engines on the *L. lactis* MG1363 peptidome. The bar chart shows the number of unique  
742 peptides identified by each search engine. The number (white) is the mean value from  
743 biological triplicates. Standard deviation is shown at the top of each bar. (D)  
744 Presence/absence map of unique peptides from panel 2C, identified by the indicated 6 search  
745 engines on the biological replicates (roman numbers).  
746

747 **Figure 3 Overview of mass spectrum results from PEAKS studio of intracellular**  
748 **peptidomes of *L. lactis* MG1363 and 6 of its peptidase mutants.** (A) Peptide spectrum  
749 matches (PSMs), identified unique peptides and proteins identified in the peptidomes of the  
750 indicated strains. The bar charts show the mean value, the exact value of which is shown in  
751 the bar, and standard deviation from biological triplicates of each strain. (B) Presence/absence  
752 map of identified unique peptides from panel 3A. (C) Venn diagram of identified unique  
753 proteins from panel 3A (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

754

755 **Figure 4 Ontology analysis of unique proteins identified in the intracellular peptidomes**  
756 **of *L. lactis* MG1363 and the indicated 6 isogenic *pep* mutants.** Correlation of each protein  
757 (outer circle on the left side of each graph) and Clusters of Orthologous Groups (COG)  
758 category is represented by lines. Abbreviations of COGs are listed in the inset. The figure was  
759 generated by R package Circlize (51).

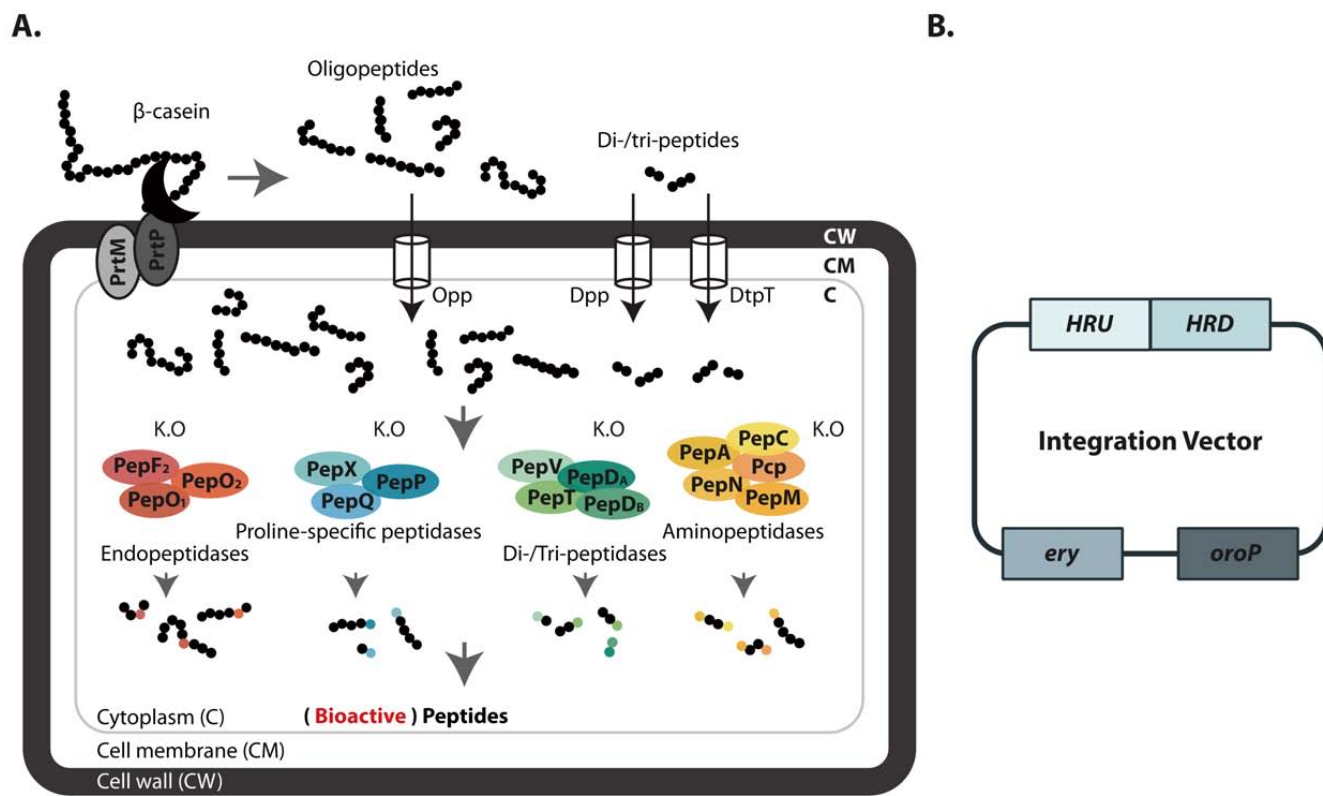
760

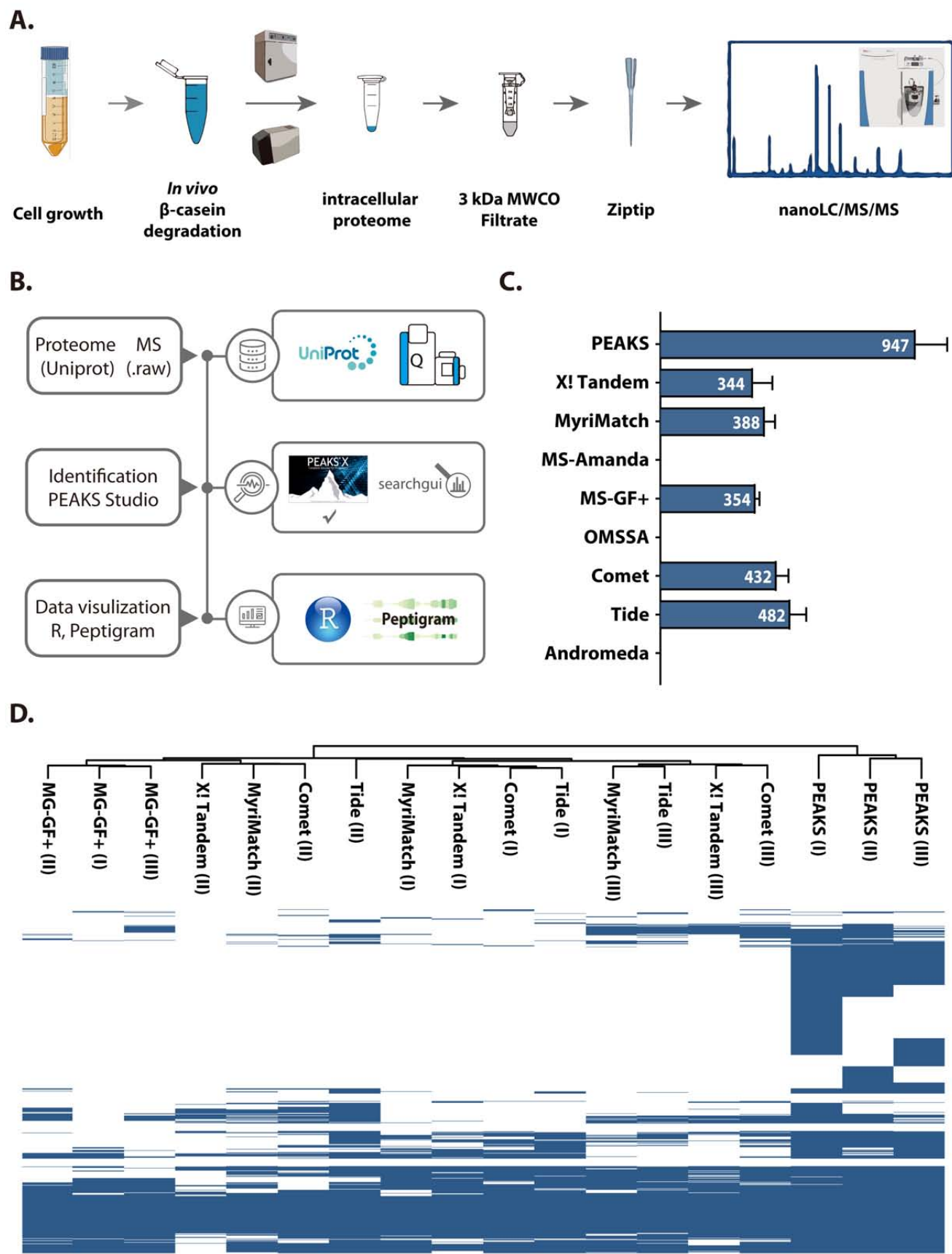
761 **Figure 5 Intracellular profile and physicochemical properties of peptides derived from**  
762  **$\beta$ -casein after its initial hydrolysis by extracellular PrtP in *L. lactis* MG1363 and its *pep***  
763 **mutants. (A)** Intracellular peptides assigned to  $\beta$ -casein by the PEAKS X studio software in the  
764 peptidomes of the indicated strains. The visualization tool Peptigram (30) was used to  
765 generate the figure. The linear sequence of bovine  $\beta$ -casein (224 amino acid residues  
766 including the signal sequence f1-15) is shown at the top. Each vertical green bar represents an  
767 amino acid residue in  $\beta$ -casein, with the height denoting the count and the color intensity  
768 corresponding to the sum of peptides overlapping at this position. (B) Violin plots with mean  
769 value (black dots) showing the physicochemical properties of the intracellular peptides  
770 assigned to  $\beta$ -casein in the peptidomes of *L. lactis* MG1363 and its peptidase mutants. The  
771 inset shows the strain identity; strains are shown in the same order in each panel.  
772

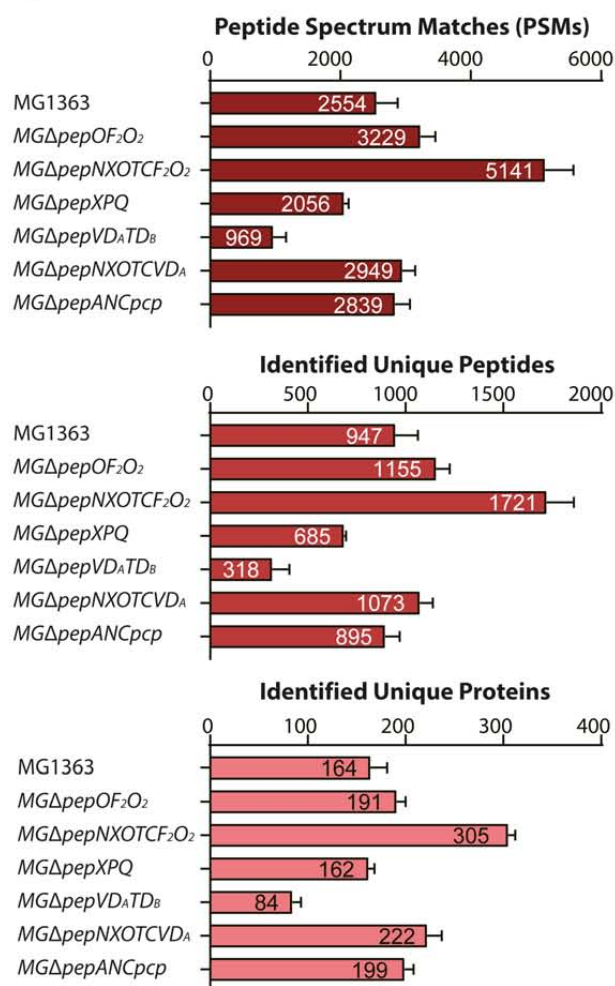
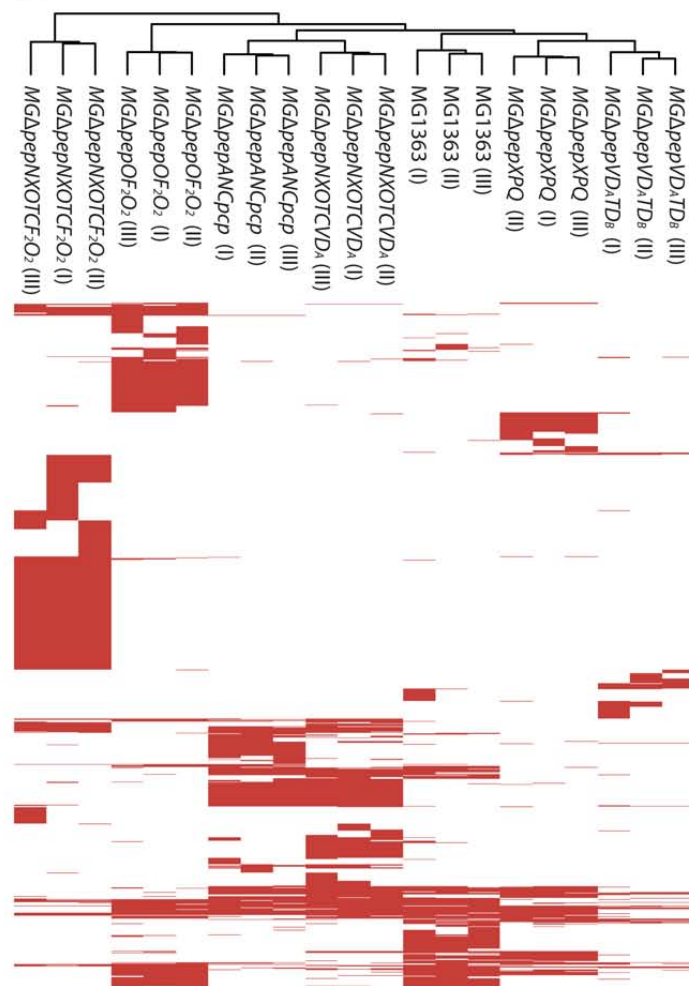
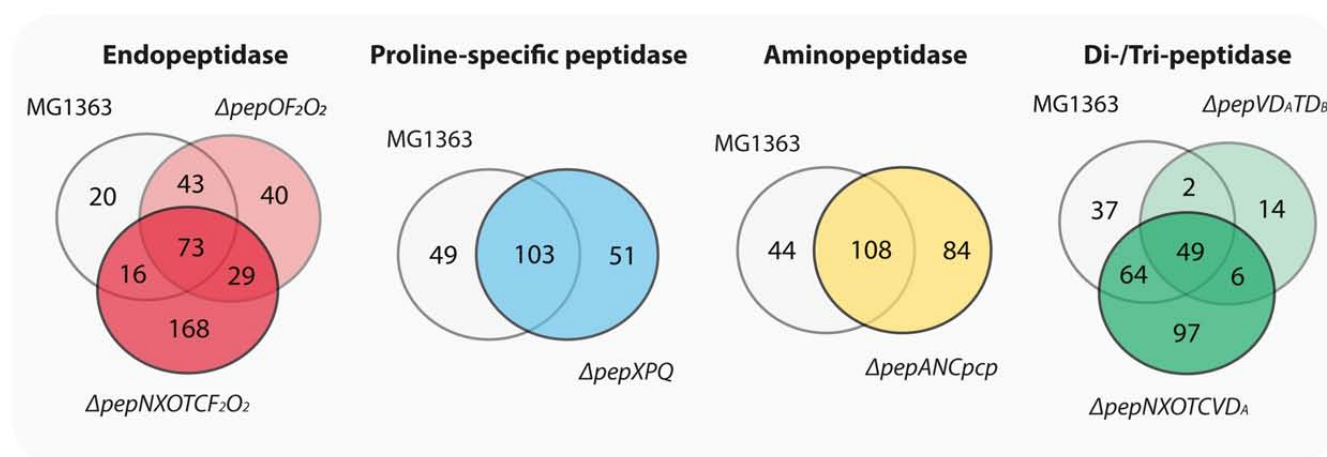


773 **Figure 6 Optimization of a  $\beta$ -casein bioactive-peptide database.** (A) Optimization strategy.  
774 BP, bioactive peptide. (B) UpSetR plot (53) highlighting the intersection of peptides with  
775 bioactivities, as indicated by the bullet points. Horizontal bars (Set size) indicate the number  
776 of bioactive peptides for each bioactivity. The vertical bars (Intersection size) show the  
777 number of peptides with a certain bioactivity activity. A peptide can have more than 1  
778 bioactivity, as indicated by the connected bullet points. For example, the set size of PEP-I is  
779 four of which 3 peptides have only PEP-I activity while 1 peptide has both PEP-I and ACE-I  
780 activity. PEP-I: prolyl endopeptidase inhibitory, AntiO: antioxidation, AntiM: antimicrobial,  
781 DPP-IV-I: dipeptidyl peptidase IV inhibitory, Immuno: immunomodulatory, ACE-I: angiotensin-  
782 converting enzyme inhibitory.  
783

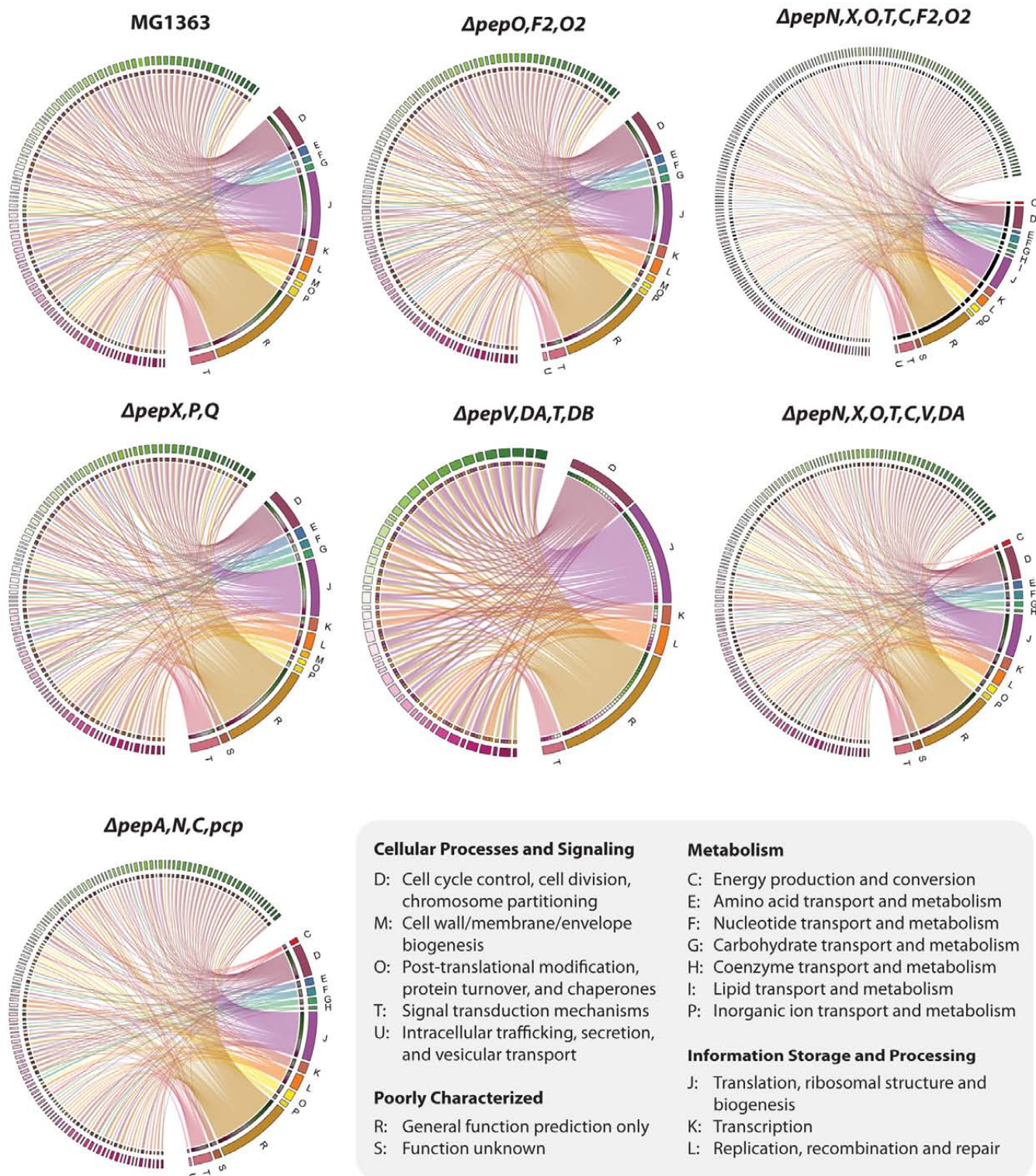
784 **Fig 7.  $\beta$ -casein-derived peptides identified in the intracellular peptidomes of**  
785 ***MG $\Delta$ pepOF<sub>2</sub>O<sub>2</sub>* (peptide sequences in red) and MG1363 (peptide sequences in black).**  
786  
787

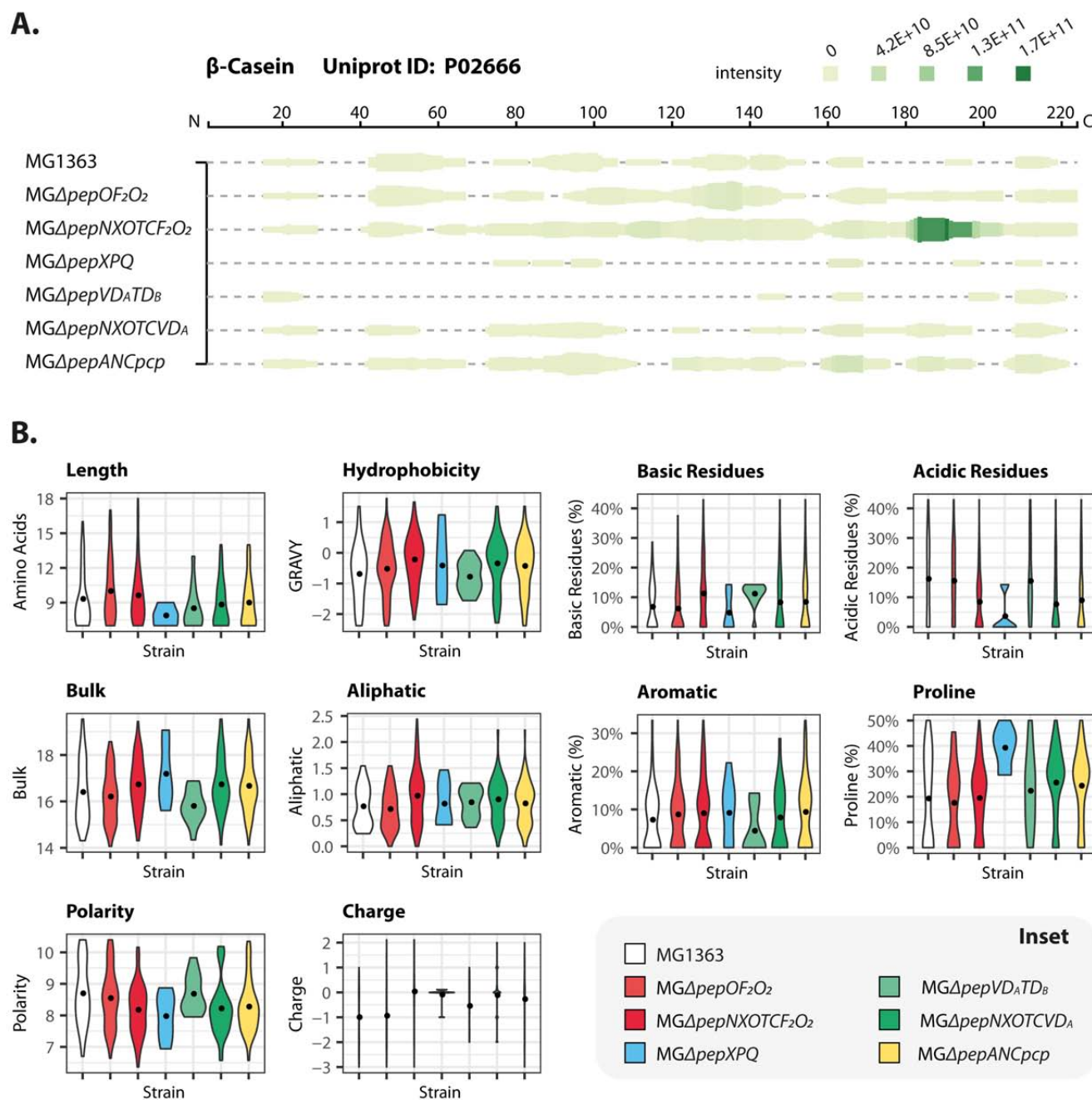


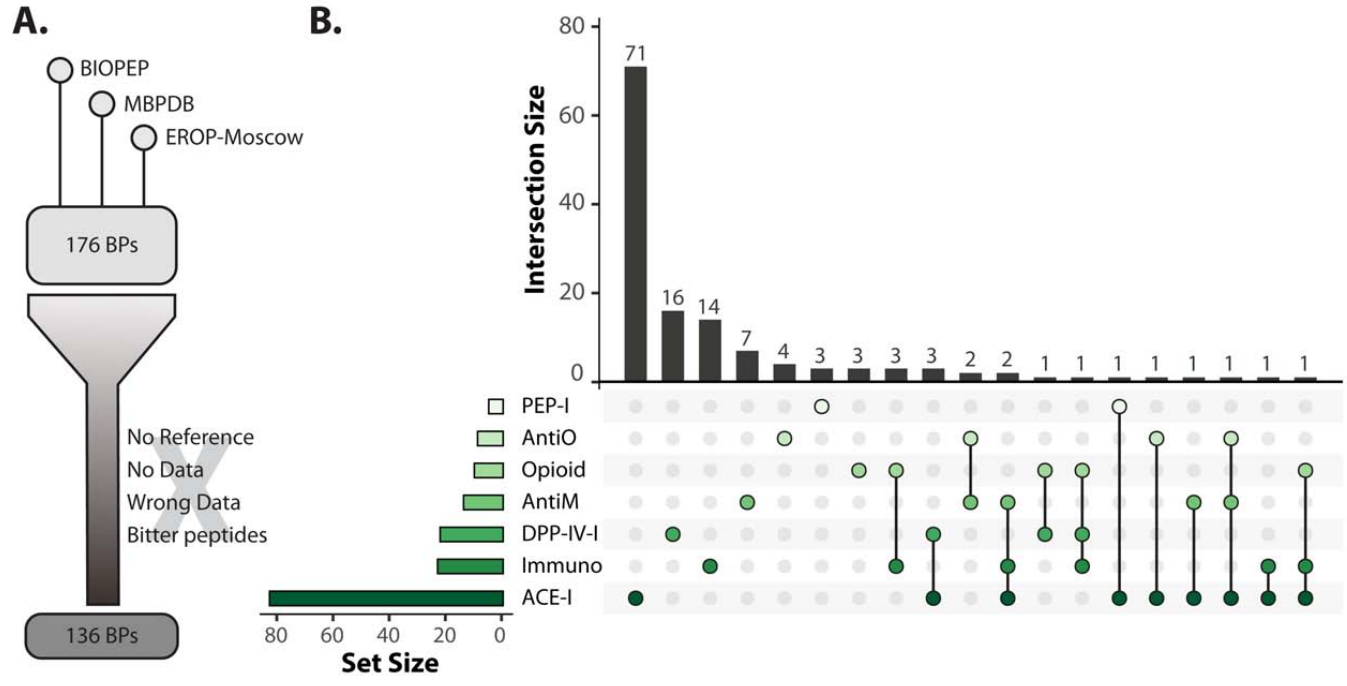


**A.****B.****C.**













Strain	Identified peptides		
 MGΔ <i>pepOF</i> <sub>2</sub> O <sub>2</sub>	<b>TQTPVVVPPFLQPEVM</b>	<b>TPVVVPPFLQPEVM</b>	<b>PVVVPPFLQPEVM</b>
 MG1363	<b>TQTPVVVPPFLQPE</b> <b>TQTPVVVPPF</b> <b>TQTPVVVPP</b>	<b>TPVVVPPF</b> <b>TPVVVPP</b>	<b>PVVVPPF</b>